

# **Predicate-induced semantic prominence in online argument linking: experiments on affectedness and analytical tools**

## **Dissertation**

zur Erlangung des akademischen Grades

**Doctor philosophiae**  
**(Dr. phil.)**

eingereicht

an der Philosophischen Fakultät II  
der Humboldt-Universität zu Berlin

von

Mag. phil. Philip Rausch

Präsident der Humboldt-Universität zu Berlin

Prof. Dr. Jan-Hendrik Olbertz

Dekanin der Philosophischen Fakultät II

Prof. Dr. Helga Schwalm

Gutachter:

1. Prof. Dr. Manfred Krifka
2. Prof. Dr. Werner Sommer
3. Prof. Jeffrey S. Morris, PhD

Datum der Verteidigung: 25.01.2016



*Für meine Familie*  
*Hex hex! Bing Bing!*



# Acknowledgements

Writing a dissertation can be a solitary process, but getting it done requires the help, support, patience and input of many people and institutions. Of these, I would first like to thank the Berlin School of Mind and Brain for funding my work. In particular, my gratitude goes to Annette Winkelmann and Anne Löchte for always taking care of the students, even in the most complex situations.

I started out my work on this thesis with two supervisors, Manfred Krifka and Werner Sommer. I am deeply grateful to them for accompanying me throughout the course of the project with their sound advice, endurance and support on all conceivable levels. While I have learned a lot about doing science from them, the lessons learned go far beyond this. While working on the second, methodological part of this thesis, Jeff Morris became my third supervisor, guiding me not only through a whole number of topics, issues and challenges whenever needed, but in addition providing support wherever possible. I would like to thank all three of them for their detailed and always helpful feedback on the various parts of this thesis – needless to say that any remaining errors are my own!

My special gratitude goes to all my friends and colleagues who have shared life, time, work, office, beer, ups and downs and much more during all this time with me! While there are too many to list, I would like to thank Mareike Bayer, Milena Rabovsky, Thomas Pinkpank and Ulrike Bunzenthall (also) for helping me with the challenges of setting up and carrying out EEG experiments as well as Felix König for a lot of things, including help with some of the formal aspects in the second part of the thesis, Leonardo Lancia for providing very valuable feedback on the packages presented there, as well as Anna Czypionka.

Finishing this thesis has taken a long time, during which I lost my uncle Stephan, who was always curious about and deeply interested in my work. Just as he would have loved to see this thesis done, I would have loved to try and answer the countless questions he would certainly have had about it and share this moment with him. It hurts that I can't.

This thesis would never have seen the light of day without my family, whose unconditional love and support have carried me through all joys and sorrows. My mum and dad have not only helped and encouraged me in every conceivable way, but have at the same time been the most amazing grandparents for our kids Nella and Luk, thus teaching me many important things about being a parent. Without all the dedication of Oma und Opa Wien, as well as that of Oma und Opa Garten (formerly also known as Birgit and Helmut), I don't know how we would have coped with the madness of ordinary family life! Equally, Tobi was not only there for me as a brother, but whenever possible for Nella and Luk, who, no doubt, will one day follow him and Nathalie up the highest mountains. Finding words for the role that my wife Marita and our kids are playing in my life is difficult. Thanks for bringing all your love, joy, energy and craziness into my life! Thanks for your endless patience, endurance and support, for the chaos and the calm! Thanks for driving me nuts and keeping me sane!



# Abstract

The present thesis consists of two parts. Part I investigates the effects of a specific semantic characteristic of syntactically transitive predicates on different aspects of online sentence processing: the degree of *affectedness* such a predicate implies for the argument realised as the direct object, which is taken to indicate the extent to which the respective event participant necessarily undergoes a change of state during the event expressed. For example, while it is clear that *the window* changes maximally in an event such as *The boy broke the window*, it may survive unscathed in *The boy hit the window*. In part I, I present three experiments (using acceptability judgements, reading times and event-related potentials – ERPs) investigating the effects of affectedness on the processing of the predicate itself as well as on the integration of following argument noun phrases (NPs), thus providing a picture of the impact of affectedness on lexical-semantic processing (as measured on the predicates) and its possible influence in online argument integration (as measured on the NP arguments) during sentence reading. Arguing that the effects of affectedness on processes related to argument linking are maximised when using German *-ung* nominalisations derived from verbs implying different degrees of affectedness for the event participant realised as direct object (such as *Bewunderung/admiration* or *Ermordung/assassination*), rather than verbs themselves, the experiments make use of a paradigm involving the linking of either subject or object genitive argument NPs to such deverbal, eventive nominalisations. The graded approach to affectedness developed by Beavers (2010, 2011) allows to view the degree of affectedness implied for an object argument as a form of semantic argument prominence induced by the predicate, providing links to prototype-based theories of argument linking and thus also to a specific model of sentence processing, the Extended Argument Dependency Model (eADM – Bornkessel & Schlesewsky, 2006), which attributes a central role to these aspects of NP arguments in sentence parsing. While no clear evidence for effects of affectedness on the predicate (the nominalisations) was found in the two online processing experiments, the analysis of the data related to the integration of either subject or object genitives converged across experiments and revealed a specific interaction between the degree of affectedness and the acceptability of linking either subject or object genitives. These acceptability patterns were accompanied by consistent reading time and ERP interaction effects. Analysis of the ERP data using the wavelet-based functional mixed model (WFMM – Morris & Carroll, 2006) suggests a prominent role of two temporally concurrent late positivities with left posterior and anterior midline foci in the integration of the genitive arguments. These findings are discussed in the context of the theoretical and sentence processing models mentioned above, considering possible roles of argument prototypicality and of differences between predicate-induced and argument-inherent semantic argument prominence (such as animacy and definiteness). The second part introduces two software packages for the statistical environment R which were used for the WFMM-based analysis of the ERP data in the first part: the *wrapfmm* package provides a high-level wrapper for the WFMM executable, while the *stepmom* (spatiotemporal electrophysiological model maps) package allows to model multi-channel ERP data using the WFMM.





# Abstract (Deutsch)

Die vorliegende Arbeit besteht aus zwei Teilen. Der erste Teil untersucht den Effekt eines bestimmten semantischen Merkmals von syntaktisch transitiven Prädikaten auf verschiedene Aspekte der Satzverarbeitung: den Grad der *Affiziertheit*, den solch ein Prädikat für das als direktes Objekt verwirklichte Argument impliziert. Dieser gibt an, in welchem Ausmaß dieser Ereignisteilnehmer während des ausgedrückten Ereignisses notwendigerweise eine Zustandsveränderung erfährt. Während beispielsweise klar ist, dass *das Fenster* in einem Ereignis wie *Der Junge zerbrach das Fenster* sich maximal verändert, könnte es ein Ereignis wie *Der Junge traf das Fenster* unbeschadet überstehen. Teil I stellt drei Experimente (zu Akzeptabilitätsurteilen, Lesezeiten und ereigniskorrelierte Potentialen –EKP) vor, die die Effekte von Affiziertheit auf die Verarbeitung des Prädikats selbst sowie auf die Integration von folgenden Argument-Nominalphrasen (NPn) mit dem Ziel untersuchen, Einblicke zur Rolle der Affiziertheit auf die lexikalisch-semantische Verarbeitung (gemessen auf den Prädikaten) und möglichen Auswirkungen auf die Argumentintegration während des Satzlesens (gemessen auf den Argument-NPn) zu gewinnen. Hierbei wird die Annahme gemacht, dass die Effekte von Affiziertheit auf Prozesse, die dem Argumentlinking zugrunde liegen, im Deutschen durch die Verwendung von deverbalen *-ung*-Nominalisierungen maximiert werden. Statt der Verben selbst wurden in den Experimenten deshalb *-ung*-Nominalisierungen verwendet, die von transitiven Verben abgeleitet werden, die unterschiedliche Grade an Affiziertheit für den als direktes Objekt verwirklichten Ereignisteilnehmer implizieren (z.B. *Bewunderung* vs. *Ermordung*). Innerhalb dieses Paradigmas wurden die Effekte auf das Argumentlinking anhand der Integration von Subjekts- und Objektsgenitivargument-NPn mit solch deverbalen, eventiven Nominalisierungen untersucht. Der von Beavers (2010, 2011) entwickelte graduelle Ansatz zu Affiziertheit erlaubt es, den für ein Objektargument implizierten Affiziertheitsgrad als eine Form der semantischen Argumentprominenz zu sehen, die vom Prädikat induziert wird. Er bietet so einerseits eine Verbindung zu Prototyp-basierten Theorien des Argumentlinkings und andererseits zu dem Extended Argument Dependency Model (eADM – Bornkessel & Schlesewsky, 2006), einem Satzverarbeitungsmodell, in dem diese Eigenschaften von Argument-NPn eine zentrale Rolle einnehmen. Bei den Akzeptabilitätsurteilen zeigte sich ein spezifisches Interaktionsmuster zwischen der Akzeptabilität eines Subjekts- oder Objektsgenitivs und dem Grad der Affiziertheit. Während sich keine klaren Effekte von Affiziertheit bei der Verarbeitung der Prädikate (Nominalisierungen) in den zwei Verarbeitungsexperimenten zeigten, ergaben sich bei der Integration von Subjekts-/Objektsgenitiven mit den Befunden zur Akzeptabilität übereinstimmende Interaktionseffekte. Die Analyse der EKP-Daten anhand des wavelet-based functional mixed model (WFMM – Morris & Carroll, 2006) legt nahe, dass hierbei zwei parallel auftretende späte Positivierungen über links-posteriores und frontozentralen Elektroden eine zentrale Rolle spielen. Diese Befunde werden im Kontext der oben erwähnten Modelle zur Affiziertheit und Satzverarbeitung diskutiert, mit Fokus auf dem möglichen Einfluss von Argumentprototypikalität und Unterschieden zwischen solch prädikatsinduzierter und argumentsinhärenter (z.B. Belebtheit oder Definitheit) semantischer Prominenz. Der zweite Teil stellt zwei Software-Pakete für die Statistik-Umgebung R vor, die für die WFMM-basierte Analyse der EKP-Daten im ersten Teil verwendet wurden: das *wrapfmm* Paket stellt eine R-Schnittstelle für die WFMM Kernsoftware dar, das *stepmom* (**s**patiotemporal **e**lectrophysiological **m**odel **m**aps) Paket erlaubt die Modellierung von Multikanal-EKP-Daten mit Hilfe des WFMM.



# Contents

List of Figures	vi
List of Tables	viii
List of Abbreviations	xi
Preface at the time of publication	xiii
Introduction	1
<b>I Affectedness as a factor in online argument linking: experiments in the nominal domain</b>	<b>9</b>
<b>1 Affectedness as a pivot between predicate meaning and syntactic structure</b>	<b>11</b>
1.1 Argument linking: some basics . . . . .	12
1.2 Grammatically relevant aspects of verb meaning . . . . .	15
1.3 Thematic roles and argument linking . . . . .	19
1.3.1 Thematic roles . . . . .	19
1.3.2 Thematic hierarchies and prominence scales . . . . .	23
1.4 Alternatives to thematic role-based approaches . . . . .	25
1.4.1 Predicate decompositions . . . . .	25
1.4.2 Prototype-based argument linking approaches . . . . .	28
1.4.2.1 Dowty's Proto-Roles . . . . .	28
1.4.2.2 Extensions and adaptations of Dowty's Proto-role approach . . . . .	32
1.5 Focussing on affectedness . . . . .	33
1.5.1 Affectedness as a graded phenomenon influencing semantic argument prominence . . . . .	34
1.5.2 Correlating phenomena . . . . .	41
1.5.2.1 Transitivity . . . . .	41
1.5.2.2 Lexical aspect . . . . .	45
1.6 Summary . . . . .	48
<b>2 Probing affectedness: a detour via argument linking in deverbal nominalisations</b>	<b>49</b>
2.1 Argument linking: from verbs to nouns . . . . .	50
2.2 Linking patterns in German deverbal <i>-ung</i> nominalisations . . . . .	53
2.3 Change of state as a semantic determinant of argument linking in eventive <i>-ung</i> nominalisations . . . . .	57

2.4	Determining linking in <i>-ung</i> nominalisations: a case of semantic prominence induced by degrees of affectedness? . . . . .	60
2.5	Excursus: on the syntactic status of postnominal genitives – arguments or adjuncts? . . . . .	65
2.6	Eventive nominalisations as a backdoor to affectedness-related linking effects at the syntax/semantics interface . . . . .	68
2.7	Acceptability judgement experiment . . . . .	71
2.7.1	Operationalising degree of affectedness as a continuous variable . . . . .	72
2.7.2	Pretest . . . . .	74
2.7.2.1	Materials . . . . .	74
2.7.2.2	Procedure . . . . .	76
2.7.2.3	Participants . . . . .	77
2.7.2.4	Results . . . . .	77
2.7.3	Acceptability judgement experiment . . . . .	79
2.7.3.1	Materials . . . . .	79
2.7.3.2	Procedure . . . . .	80
2.7.3.3	Hypotheses . . . . .	81
2.7.3.4	Participants . . . . .	82
2.7.3.5	Data analysis . . . . .	83
2.7.3.6	Results . . . . .	84
2.7.3.7	Discussion . . . . .	86
2.8	Summary . . . . .	92
<b>3</b>	<b>Psycholinguistic backgrounds</b>	<b>93</b>
3.1	From offline judgements to online processing: aims . . . . .	94
3.2	Psycholinguistic backgrounds . . . . .	94
3.2.1	Processing correlates of verbal lexical-semantic complexity . . . . .	94
3.2.2	Interactions of verbal semantics and syntactic processing . . . . .	100
3.2.3	ERP-correlates of conceptual-semantic and syntactic processing: traditional dichotomies and more recent findings . . . . .	101
3.2.4	The Extended Argument Dependency Model and the processing of core predicate-argument relations . . . . .	106
3.2.4.1	Scope, architecture and principles . . . . .	107
3.2.4.2	Relevant ERP-markers and their functional interpretations . . . . .	110
3.2.4.3	The eADM and the present work: common ground and differences in perspective . . . . .	113
3.2.5	Similarity-based interference in sentence processing . . . . .	116
3.2.6	Relevant aspects of discourse processing . . . . .	118
3.3	Summary . . . . .	120
<b>4</b>	<b>Affectedness as a factor in online argument linking: reading time and electrophysiological evidence</b>	<b>121</b>
4.1	Self-paced reading experiment . . . . .	121
4.1.1	Materials . . . . .	122
4.1.2	Procedure . . . . .	124
4.1.3	Hypotheses . . . . .	125
4.1.4	Participants . . . . .	126
4.1.5	Data analysis . . . . .	126

4.1.6	Results	128
4.1.7	Discussion	137
4.1.8	Summary	146
4.2	ERP experiment	147
4.2.1	Materials	147
4.2.2	Procedure	148
4.2.3	EEG recording and preprocessing	148
4.2.4	Hypotheses	149
4.2.5	Participants	151
4.2.6	Data analysis	151
4.2.7	Results	158
4.2.8	Discussion	170
4.3	Summary	184
<b>5</b>	<b>General discussion</b>	<b>187</b>
5.1	Generalisability	187
5.2	Research question one	190
5.3	Research question two	191
5.4	Research question three	193
5.5	Other issues	196

## II Got curves?

### Introducing two R-packages for WFMM-based analysis of curve- and ERP-data 199

<b>6</b>	<b>Backgrounds: towards holistic analyses of curve data</b>	<b>201</b>
6.1	Limitations of the traditional ANOVA approach to the analysis of scalar data	202
6.1.1	The challenge: the 'language-as-a-fixed-effect-fallacy'	202
6.1.2	The ANOVA fix: $F'/F'_{min}$ and the $F_1 \times F_2$ criterion	203
6.1.3	Problems with ANOVA coping strategies: matching, factorisation, aggregation and correction	204
6.2	Linear mixed models as an alternative analysis tool	206
6.2.1	Flexibility	207
6.2.2	Open issues	208
6.2.3	Design structure	209
6.3	Challenges of analysing curve data: towards holistic approaches	209
6.3.1	Holistic analyses: capturing design and object structure	209
6.3.2	Standard approaches: failing holism (one way or the other)	210
6.4	Challenges of ERP data analysis	212
6.4.1	ERP data: standard analyses approaches and their shortcomings	212
6.4.2	Alternative approaches	213
6.5	An alternative framework for the analysis of curve data: functional data analysis	216
6.5.1	The general framework	216
6.5.2	Functional mixed effects models	219
6.6	The wavelet-based functional mixed model	222
6.6.1	Model definition	222
6.6.2	Wavelets: a quick tour	223

6.6.3	Internal WFMM architecture, modelling processes and scope . . . . .	227
6.6.4	Inference in the WFMM . . . . .	230
6.6.5	The software: features and limitations . . . . .	232
<b>7</b>	<b>The <i>wrapfmm</i> package</b>	<b>235</b>
7.1	Aims and scope . . . . .	235
7.2	Availability & compatibility . . . . .	236
7.3	Implementation . . . . .	236
7.4	WFMM backend . . . . .	236
7.5	Limitations . . . . .	237
7.6	Overview . . . . .	238
7.6.1	Getting oriented . . . . .	238
7.6.2	A quick tour . . . . .	239
7.6.2.1	The data set . . . . .	239
7.6.2.2	<code>wfmm_base</code> : setting up a base object . . . . .	240
7.6.2.3	<code>wfmm</code> : running a wavelet-based functional mixed model . . . . .	240
7.6.2.4	<code>summary</code> : summarising a WFMM's posterior estimates . . . . .	241
7.6.2.5	<code>follow_up</code> : computing post-hoc posteriors . . . . .	244
7.6.2.6	<code>plot</code> : plotting data and/or fixed effects posteriors . . . . .	246
7.7	Posterior inference: computation of credible intervals and posterior probabilities . . . . .	250
7.7.1	Credible intervals . . . . .	250
7.7.1.1	Pointwise credible intervals . . . . .	250
7.7.1.2	Simultaneous credible bands . . . . .	250
7.7.2	Posterior probabilities . . . . .	251
7.7.2.1	Pointwise posterior probabilities . . . . .	252
7.7.2.2	'Simultaneous posterior probabilities' . . . . .	252
7.8	Summary . . . . .	260
<b>8</b>	<b>The <i>stepmom</i> package</b>	<b>261</b>
8.1	Aims and scope . . . . .	261
8.2	Availability & compatibility . . . . .	262
8.3	Implementation . . . . .	262
8.4	WFMM backend . . . . .	262
8.5	Limitations . . . . .	262
8.6	Overview . . . . .	263
8.6.1	Getting oriented . . . . .	263
8.6.2	A quick tour . . . . .	265
8.6.2.1	The data set . . . . .	265
8.6.2.2	<code>mom_base</code> : setting up a base object . . . . .	265
8.6.2.3	<code>vom</code> : plotting ERP data as voltage scalp maps . . . . .	265
8.6.2.4	<code>wfmm</code> : fitting wavelet-based functional mixed models to multi-channel ERP data . . . . .	269
8.6.2.5	<code>mom</code> : summarising a WFMM's posterior estimates as 'model maps' . . . . .	270
8.6.2.6	<code>ani</code> : animated voltage or model maps . . . . .	273
8.6.2.7	<code>summary</code> and <code>plot</code> : summarising a WFMM's posterior estimates and plotting the underlying data for selected electrodes . . . . .	274
8.6.2.8	<code>follow_up</code> : computing post-hoc posteriors . . . . .	275

8.6.2.9 lmerMom: fitting LMMs to time window average data . . . .	276
8.7 Summary . . . . .	280
<b>9 General discussion</b>	<b>281</b>
9.1 Summary . . . . .	281
9.2 Discussion . . . . .	282
9.2.1 Meeting and failing holism with the current approach in practice . . . .	282
9.2.2 Possible extensions for modelling EEG data with the WFMM . . . . .	284
9.2.3 Improving information transmission . . . . .	286
9.3 Conclusion . . . . .	286
<b>Bibliography</b>	<b>289</b>
<b>Appendix A Nominalisation infos</b>	<b>309</b>
<b>Appendix B Supplementary materials for the acceptability rating study</b>	<b>313</b>
<b>Appendix C Supplementary materials for the self-paced reading study</b>	<b>325</b>
<b>Appendix D Supplementary materials for the event-related potentials study</b>	<b>339</b>

# List of Figures

2.1	AI rating from pretest for each verb. . . . .	78
2.2	By-subject and by-nominalisation acceptance proportions for each 'Linking' condition. . . . .	85
2.3	Predictions derived from the GLMM for the acceptability judgement data together with 95% bootstrap CIs. . . . .	87
3.1	Architecture of the eADM as presented by Bornkessel-Schlesewsky and Schlesewsky (2009). . . . .	108
4.1	Self-paced reading rating data: by-subject and by-nominalisation mean acceptability ratings for each 'Linking' level. . . . .	129
4.2	LMM-based predictions for the acceptability rating data of the self-paced reading experiment. . . . .	131
4.3	Reading times for six regions. . . . .	133
4.4	LMM-based predictions for RTs in the 'PostGenitive 1' region. . . . .	138
4.5	Non-linear RT smooths for two regions. . . . .	140
4.6	Toy model of the SPR results. . . . .	144
4.7	ERP rating data: by-subject and by-nominalisation mean acceptability ratings for each 'Linking' level. . . . .	158
4.8	LMM-based predictions for the acceptability rating data of the ERP experiment. . . . .	160
4.9	Nominalisation segment: model maps for the 's.AI' predictor. . . . .	164
4.10	Genitive segment: Model maps for the 'c.Linkin9' predictor. . . . .	165
4.11	Genitive segment: model maps for the 's.AI' predictor. . . . .	167
4.12	Genitive segment: Model maps for the 'c.Linkin9:s.AI' predictor. . . . .	168
4.13	Genitive segment: ERP interaction patterns at electrodes Fz and P5. . . . .	169
4.14	Genitive segment: Follow-up model maps for simple slopes of 's.AI'. . . . .	171
4.15	Genitive segment: Follow-up model maps for simple slope of 's.AI' in the 'GenS' condition in the very late time window. . . . .	172
4.16	Genitive segment: Linear smooths for 'AI' at each level of 'Linkin9' computed for electrodes Fz, P5 and the average across these. . . . .	176
6.1	Scaling and wavelet functions of two wavelets. . . . .	223
6.2	Simulated curve and the wavelet coefficients for its DWT. . . . .	226
6.3	Thresholded DWT coefficients plus original and smoothed curve. . . . .	227
6.4	Original and smoothed curve with lenient compression. . . . .	228
6.5	WFMM-based inference about amplitude and latency effects. . . . .	231



7.1	Posterior fixed effects summary plot of the example WFMM. . . . .	242
7.2	Posterior fixed effects summary plot for two effects. . . . .	243
7.3	Summary of posterior mean estimates of the WFMM's variance components. . . . .	244
7.4	Summary of the posterior by-subject random effects estimates. . . . .	245
7.5	Summary plot of the posterior cell means follow-up estimates. . . . .	246
7.6	Mean curves for low and high predictability words, faceted by 'Session'. . . . .	247
7.7	Same as in previous figure, with difference curves added . . . . .	247
7.8	Difference curves. . . . .	248
7.9	Difference curve compared to the posterior estimate for the Predictability fixed effect. . . . .	249
7.10	Mean curves for low and high predictability words together with posterior follow-up estimates. . . . .	250
7.11	Illustration of BOND. . . . .	255
7.12	Two types of SCBs and simultaneous posterior probabilities. . . . .	256
7.13	Comparison of BONDS for different CI/SCB levels. . . . .	258
8.1	Voltage maps for three time windows produced with the vom function. . . . .	266
8.2	Voltage maps with alternative map colours and settings. . . . .	266
8.3	Voltage maps for each level of the 'Predictability' factor. . . . .	267
8.4	Difference maps. . . . .	268
8.5	Difference maps for two subjects, split by 'Session'. . . . .	269
8.6	Fixed effects model maps. . . . .	271
8.7	Model maps for posterior mean estimates and posterior probabilities. . . . .	272
8.8	Variance maps. . . . .	273
8.9	Posterior fixed effects summary for two electrodes. . . . .	275
8.10	Difference curves with posterior estimates. . . . .	276
8.11	Model maps of the posterior cell means estimates for low and high predictab- ility words. . . . .	277
8.12	Model maps of the lmerMom-based fixed effects estimates. . . . .	279
8.13	Variance maps of the lmerMom-based model. . . . .	279
D.1	Genitive segment: lmer-based model maps for the 'c.Linking' predictor. . . . .	341
D.2	Genitive segment: lmer-based model maps for the 's.AI' predictor. . . . .	342
D.3	Genitive segment: lmer-based model maps for the 'c.Linking:s.AI' interaction predictor. . . . .	342

# List of Tables

1.1	Proto-role properties of Dowty (1991, p. 572). . . . .	29
1.2	Grimm's (2011, p. 520) privative Agentivity features defining prototypical Agents. . . . .	33
1.3	Verbal predicates grouped into four different affectedness classes via a subset relation following Beavers (2011). . . . .	36
1.4	Possible correspondences between the Proto-Patient entailments originally proposed by Dowty (1991) and those of Beavers (2010). . . . .	40
1.5	Transitivity components of Hopper and Thompson (1980, p. 252). . . . .	42
2.1	Random and fixed effects summary tables for the GLMM for the acceptability judgement data. . . . .	86
2.2	Random and fixed effects summary tables for the GLMM for the GenO subset of the acceptability judgement data. . . . .	86
2.3	Random and fixed effects summary tables for the GLMM for the GenS subset of the acceptability judgement data. . . . .	88
2.4	Auxiliary Selection Hierarchy of Sorace (2000, p. 863) . . . . .	91
4.1	Random and fixed effects summary tables for the LMM for the acceptability rating data. . . . .	130
4.2	Random and fixed effects summary tables for the LMM for the GenO subset of the acceptability rating data. . . . .	132
4.3	Random and fixed effects summary tables for the LMM for the GenS subset of the acceptability rating data. . . . .	132
4.4	Random and fixed effects summary tables for the LMM results for the RT-data of the 'Nominalisation' region. . . . .	134
4.5	Random and fixed effects summary tables for the LMM results for the RT-data of the 'Genitive Article' region. . . . .	135
4.6	Random and fixed effects summary tables for the LMM results for the RT-data of the 'Genitive' region. . . . .	136
4.7	Random and fixed effects summary tables for the LMM results for the RT-data of the 'PostGenitive 1' region. . . . .	137
4.8	Random and fixed effects summary tables for the LMM results for the RT analysis of the 'GenO' subset of the 'PostGenitive 1' region. . . . .	139
4.9	Random and fixed effects summary tables for the LMM results for the RT analysis of the 'GenS' subset of the 'PostGenitive 1' region. . . . .	140
4.10	Random and fixed effects summary tables for the LMM for the ERP acceptability rating data. . . . .	159

4.11 Random and fixed effects summary tables for the LMM for the GenO subset of the ERP acceptability rating data. . . . .	161
4.12 Random and fixed effects summary tables for the LMM for the GenS subset of the ERP acceptability rating data. . . . .	161
4.13 Information about the effect clusters associated with the 's.AI' predictor for the nominalisation segment. . . . .	162
4.14 Information about the effect clusters associated with the three interest predictors for the genitive segment. . . . .	163
4.15 Information about effect clusters associated with the simple slopes of 's.AI' for the genitive segment. . . . .	169
7.1 Details for the effect clusters for electrode FCz, based upon four different inference bases. . . . .	254
7.2 'Internal' Sensitivity and specificity of different posterior probability types in the complete 63-electrode model of the ERP experiment. . . . .	257
7.3 Details for effect clusters based on BONDS-inference for three different CI-coverage levels and varying significance thresholds. . . . .	259
A.1 Infos on nominalisations used in the three experiments. . . . .	309
C.1 Random and fixed effects summary tables for the CLMM for the SPR acceptability rating data. . . . .	325
C.2 Random and fixed effects summary tables for the CLMM for the GenO subset of the acceptability rating data. . . . .	326
C.3 Random and fixed effects summary tables for the CLMM for the GenS subset of the acceptability rating data. . . . .	326
D.1 Random and fixed effects summary tables for the CLMM for the ERP acceptability rating data. . . . .	339
D.2 Random and fixed effects summary tables for the CLMM for the GenO subset of the acceptability rating data. . . . .	340
D.3 Random and fixed effects summary tables for the CLMM for the GenS subset of the acceptability rating data. . . . .	340



# List of Abbreviations

<b>A</b>	Adjective	<b>FDA</b>	Functional data analysis
<b>ACC</b>	Accusative	<b>FDR</b>	False discovery rate
<b>ACC</b>	Anterior cingulate cortex	<b>FEM</b>	Feminine
<b>AI</b>	Affectedness Index	<b>FMM</b>	Functional mixed effects model
<b>AMF</b>	Anterior midline field	<b>fMRI</b>	Functional magnetic resonance imaging
<b>ANOVA</b>	Analysis of variance		
<b>BOND</b>	Band of no difference	<b>GAM</b>	Generalised additive model
<b>CAR</b>	Conditional autoregressive model	<b>GEN</b>	Genitive
<b>CEN</b>	Complex event nominal	<b>GenO</b>	Object genitive (genitivus objectivus)
<b>CI</b>	Confidence/credible interval	<b>GenS</b>	Subject genitive (genitivus subjectivus)
<b>CTVs</b>	Core transitive verbs	<b>GLM</b>	General linear model
<b>CWT</b>	Continuous wavelet transform	<b>GR</b>	Generalised semantic role
<b>DAT</b>	Dative	<b>Hz</b>	Hertz
<b>DWT</b>	Discrete wavelet transform	<b>IDWT</b>	Inverse discrete wavelet transform
<b>eADM</b>	Extended Argument Dependency Model	<b>ISP</b>	Inherent suppression potential
<b>EEG</b>	Electroencephalography	<b>LAN</b>	Left anterior negativity
<b>ELAN</b>	Early left anterior negativity	<b>LIFG</b>	Left inferior frontal gyrus
<b>ERP</b>	Event-related potential	<b>LMM</b>	Linear mixed effects model
		<b>LS</b>	Logical structure
<b>ERRC</b>	Event-related regression coefficients	<b>LSS</b>	Lexical-semantic structure
<b>ESR</b>	External suppression resources	<b>MAP</b>	Morphosyntactic Alignment Principle

<b>MASC</b>	Masculine	<b>PST</b>	Past
<b>MCMC</b>	Markov chain Monte Carlo	<b>pwCI</b>	Pointwise credible intervals
<b>MEG</b>	Magnetoencephalography	<b>R-WFMM</b>	Robust wavelet-based functional mixed model
<b>MTG</b>	Middle temporal gyrus	<b>RN</b>	Result nominal
<b>MTS</b>	Meaning through syntax	<b>ROI</b>	Region of interest
<b>MULTIBONDS</b>	Multiplicity-induced band of no difference scores	<b>ROPE</b>	Region of practical equivalence
<b>N</b>	Noun	<b>RT</b>	Reading time
<b>NCTVs</b>	Non-core transitive verbs	<b>SCB</b>	Simultaneous credible band
<b>NOM</b>	Nominative	<b>SEN</b>	Simple event nominal
<b>NP</b>	Noun phrase	<b>SG</b>	Singular
<b>Nref</b>	Referentially induced frontal negativity	<b>sim.p</b>	Simultaneous probabilities
<b>OT</b>	Optimality Theory	<b>SIMBAS</b>	Simultaneous band scores
<b>P</b>	Preposition	<b>SOA</b>	Stimulus onset asynchrony
<b>PFC</b>	Prefrontal cortex	<b>SPR</b>	Self-paced reading
<b>PL</b>	Plural	<b>TSP</b>	Total suppression potential
<b>pMTG</b>	Posterior middle temporal gyrus	<b>V</b>	Verb
<b>PNP</b>	Post-N400 positivity	<b>VLPFC</b>	Ventrolateral prefrontal cortex
<b>PP</b>	Prepositional phrase	<b>WFMM</b>	Wavelet-based functional mixed model

## Preface at the time of publication

### On the up-to-dateness of the thesis and software packages presented

Quite some time has passed since the defense of this thesis (2016) and its publication (2018). Given this time gap, this preface is intended to put some aspects of this work into context. In part II, the present thesis introduces two software packages for the open statistical environment R, which were used for the analysis of the event-related potentials (ERP) data summarized in the first part (see section 4.2.6) and intended to be made available for general use. Both of them provide R-interfaces for Morris and Carroll's (2006) wavelet-based functional mixed model (WFMM) software. Since software platforms and packages move fast, this unfortunately entails that the package versions as presented here are not compatible with current versions of R and other R packages they depend on, and they interface with an old version of the WFMM software.

While I am planning to still make the packages available in the versions introduced here as supplementary materials to this thesis in a way that should allow interested readers to get them installed without serious hassle, these will in fact mainly serve the purpose of accompanying the second part of this thesis, providing the source code used and the possibility to run function calls as presented in the package introduction chapters 7 and 8<sup>1</sup>. However, I do not recommend readers who are interested in applying the approach presented to analyse their own data to make use of these package versions, one reason being that the packages as presented are not maintained.

Nevertheless, there are some good news that will hopefully compensate for this: we have been working on a number of new packages in collaboration with Prof. Jeffrey Morris, which are under active development and will be released successively, starting in the near future (check at [github.com](https://github.com)). The functionality of the *wrapfmm* R package presented in chapter 7, which provides a general user-friendly interface to the WFMM core software, will be provided by the *BayesFMM* package, which will be based on *much* improved code and will

---

<sup>1</sup>Should you be interested in the package versions as presented here, but they are not available yet, or if you have questions about the other packages mentioned here, please contact me at *rausch2018predicate* 'youknowwhat' *gmail.com* or some other address available online.

offer a lot of additional features. Thus, it will offer the possibility to not only model 1d-curves, but also 2d-images (e.g. spectrograms) or 3d-volumes as well as the possibility to process big out-of-memory data and model posteriors in a memory-friendly fashion. Beyond this, it will allow to fit robust functional mixed models (Zhu, Brown, & Morris, 2011), spatially correlated ones (Zhu, Versace, Cinciripini, Rausch, & Morris, accepted) and will incorporate numerous additional innovations. Also, other packages are under development which are intended to provide the functionality of the *stepmom* package outlined in chapter 8, which can be used to model ERP data using the WFMM. These should also offer a more complete approach to modelling ERP data within the WFMM framework than presented and applied here.

Despite the rather limited practical relevance of the packages as presented here, I hope that the second part of this thesis still provides a useful overview and discussion of standard and alternative approaches to the analysis of complex curve and ERP data as well as of the WFMM framework (see the overview in chapter 6). Importantly, it also provides details on and discussion of many aspects of the (non-standard) analysis of the ERP data presented in part I (section 4.2.6). Also, the two package overview chapters still provide a good overview of the general idea of the approach – since much of the functionality and many aspects of the user interface will be retained in the new packages, they may still provide a good complementary reading.

Of course, the study of language and language processing moves forward quickly too, just as methodological development does, and the text itself would doubtlessly have profitted from some updates, taking into account more recent findings and innovations in all these areas – I apologise for anything I may thus have missed or that may be lacking. Nevertheless, I hope that this thesis still makes for an interesting read.



# Introduction

This thesis consists of two parts: the first one reviews a series of psycholinguistic experiments on the impact of semantic properties of verbs (or predicates derived from verbs) on their own processing as well as on the integration of argument noun phrases (NPs) in sentence reading; the second one presents two software packages for the open-source statistical environment R (R Development Core Team, 2013). These were used for the analysis of the event-related potential data presented in the first part and act as interface software for Morris and Carroll's (2006) wavelet-based functional mixed model (WFMM) and provide functions for the exploration and analysis of general curve and electroencephalographic (EEG) data. While this constellation may be somewhat unusual and may appear rather arbitrary at a first glance, I hope to show in the remainder of this introductory chapter that they actually complement each other, despite the obvious thematic gap. Nevertheless, readers who are only interested in only one of these two parts need not worry, since either can also be read without delving into the other. In either case, please refer to the remarks in the preface putting this work into context at the time of publication.

## Part 1

At the heart of the first part lies the question whether and how a special facet of the meaning of verbs impacts different aspects of sentence processing: while verbs – and other content words in general – carry and convey numerous aspects of lexical meaning, only a small number of these is usually taken to be grammatically relevant, in the sense that these specific lexical meaning features are 'visible' and relevant for other levels of the linguistic system and thus play a role in regulating morphosyntactic processes. In the case of verbs, these aspects of verb meaning are often thought to determine the mapping from the lexical 'semantic structure' of a verb – in which relevant aspects of the event or situation a verb refers to are systematically organised and encoded – to the sentence level. Crucially, such representations comprise the 'event participants' implied by a verb, which are represented as arguments of the verb. When these are realised within sentences, they must somehow be mapped to grammatical functions such as subject or object, and a central question in linguistic research on the architecture of the syntax-semantics interface has been which semantic features of verbs impact this mapping – or 'linking' – process in what way (see Levin & Rappaport Hovav,

2005, for a comprehensive overview).

To get a little more concrete and to illustrate the particular semantic property of verbs which I will focus on in this thesis, I will provide a first example right away. Let's imagine a scene in which a boy kicks a rock, the rock hits a window and shatters it (following Levin & Rappaport Hovav, 2005, p. 1). Natural language offers us a myriad of different possibilities to communicate this event to others, such as the following two English sentences:

- (1) a. The boy broke the window with a rock.
- b. The boy hit the window with a rock.

From both of these sentences, we can easily identify three participants in this event: *the boy*, who is encoded as the (preverbal) subject, *the window* realised as the postverbal direct object and *the rock*, which appears within a prepositional phrase (PP) headed by *with*. Alternatively, with the verb *break*, we may reduce the number of participants to a single one as in example (2a), where *the window* appears as the subject. Trying the same with *hit*, as in (2b), however, leads to unacceptability<sup>2</sup>.

- (2) a. The window broke.
- b. \* The window hit.

Crucially, there are related English verbs which behave alike in this respect to either *break* or *hit*: thus, just like *break*, all the verbs in (3a) also allow the direct object *the glasses* to appear as a subject (3b), but those in (4) group with *hit* in not allowing their object to appear as the sole argument realised as subject (4b).

- (3) a. The boy bent/folded/shattered/cracked the glasses.
- b. The glasses bent/folded/shattered/cracked.
- (4) a. The boy slapped/struck/bumped/stroked the doll.
- b. \* The doll slapped/struck/bumped/stroked.

Fillmore (1970, p. 130) observed that the verbs in (3) all share a common semantic property with *break* in implying that the participant encoded as the direct object (*the window/the glasses*) necessarily undergoes a *change of state* in the course of the event expressed; clearly, if it is true that somebody broke a window or bent glasses, then something is different about the window and the glasses after the breaking/bending. The verbs in (4) as well as *hit*, on the other hand, do not entail such an obligatory change of state for their direct object arguments – after all, if a window gets hit and a doll struck, neither is implied to definitely undergo a change of state and they may very well survive the event unscathed. Thus, a sentence such as (5a), in which any consequences of the *hitting* event for the *door* are denied,

---

<sup>2</sup>Following linguistic conventions, I will use an asterisk (\*) to indicate grammatical unacceptability of sentences. Question marks (?) indicate relative unacceptability. The hash mark (#) indicates pragmatic infelicity or oddness due to conceptual-semantic aspects of a sentence.

is perfectly sensible, whereas the same is not possible with *shatter* (5b). Fillmore referred to the former class of verbs as 'change-of-state' verbs and to the latter one as 'surface-contact' verbs.

- (5) a. The bullets hit the door, but it didn't have a single scratch.  
b. # The bullets shattered the door, but it didn't have a single scratch.

Thus, we are faced with two classes of verbs which differ semantically in implying a definite or only a potential change of state for the direct object and which also differ with respect to licensing the syntactic process of 'shifting' the argument usually realised as direct object to subject position, opening the possibility that the semantic notion of change of state plays a role in its regulation. In fact, change of state has been identified as one of the key verb-induced semantic traits of arguments which are operative at the interface between verbal lexical semantics and syntax, and it has often been conceptualised in terms of *affectedness* of an event participant (see Beavers, 2010, 2011; Hopper & Thompson, 1980; Næss, 2004; Tsunoda, 1985, for example, and section 1.5).

Thus, *affectedness* is a semantic characteristic of arguments/event participants which is not inherent to the arguments, but rather has its source in the semantics of the verb an argument NP combines with. Importantly, it acts as one of the key components of the syntax-semantics interface in the regulation of argument realisation processes, which map the semantic core of the verb-argument complex onto sentence structure – and while it is not the only one, it will be the focus of the present thesis.

While the role of *affectedness*/change of state as a semantic determinant of argument mapping processes has been at the centre of a considerable body of work in theoretical linguistics (for a comprehensive review of factors involved in and theories of argument realisation, see Levin & Rappaport Hovav, 2005), there is a conspicuous scarcity of studies about its potential impact on the process of argument integration and related operations during sentence processing. As a result, it is currently not clear whether such grammatically relevant semantic properties of verbs and their accompanying argument NPs should receive more attention in models of sentence processing and which processing markers correlate with relevant phenomena.

The main aim of the current study is thus to reveal possible effects of *affectedness* during (visual) sentence processing with behavioural methods (acceptability judgements, self-paced reading) and electroencephalography (EEG), with the eventual goal of tracking electrophysiological markers of the processes involved and discussing possible implications for models of sentence processing. As already pointed out, *affectedness* is a semantic characteristic of argument NPs which is not inherent to them, but rather 'inherited' from the verb they compose with – there is thus an interplay of lexical-semantic properties of predicates on the one hand, and the realisation/integration of argument NPs on the other hand, and the most important questions posed by the current work accordingly are:

1. Does *affectedness* have any consequences for the processing of predicates with verbal

semantic structure? If so, can we isolate behavioural and electrophysiological processing correlates of this semantic property on the predicate itself?

2. Does predicate-induced affectedness have any impact on the integration and licensing of NP arguments following the predicate within a sentence? If so, can we identify behavioural and electrophysiological processing correlates of such processes on the respective argument NP?
3. How do the findings fit with current models of sentence processing which address these aspects from related perspectives?

As will become clear in the course of the first chapter, affectedness is closely tied to and correlates with a number of other linguistic phenomena, such as the notion of semantic transitivity and phenomena related to lexical aspect, and it is not always straightforward to keep these apart. On top of this, I will discuss the role of affectedness in prototype-based approaches to argument linking (e.g. Ackerman & Moore, 2001; Dowty, 1991), which provide a valuable additional perspective onto the present investigation. While these additional aspects incur a considerable amount of extra complexity into the matter, they also provide valuable points of contact with previous research on related aspects of sentence processing.

One model of sentence processing which offers a number of such points of contact with the current work is the Extended Argument Dependency Model (eADM) of Bornkessel and Schlesewsky (2006), which aims at identifying processes underlying the establishment of relations behind the verb-argument core complex in sentence parsing, thus providing a natural starting point for the current investigation. Importantly, an aspect the eADM highlights as one of the crucial factors in the integration of NP arguments of a verb is its prototypicality, a notion which I just pointed out to be of relevance for the present work too; however, in its presentation of the impact of argument prototypicality, the eADM largely focusses upon semantic properties of NP arguments which are *inherent* to the arguments, rather than induced by a verbal predicate. Such inherent prototypicality features include animacy (e.g. *the boy* vs. *the rock*) or definiteness (e.g. *the boy* vs. *a boy*), which are 'set' locally within the NP itself. Thus, while there is considerable thematic overlap and common ground, the perspective onto argument integration processes (see question 2 above) taken within this thesis is a complementary one, by foregrounding the role of predicate-dependent semantic characteristics of arguments.

Within the eADM, such NP-features contribute to an argument's *prominence* (see Bornkessel-Schlesewsky & Schlesewsky, 2009, especially), which is another key notion of the model intersecting with the present work. The theoretical model of the role of affectedness in argument linking processes I will use as starting point for the experimental investigations is the one of Beavers (2010, 2011), who develops a *graded* model of affectedness in which affectedness is matter of degree. Crucially, Beavers argues that his approach to affectedness is compatible with prototype-based theories of argument linking such as that of Dowty (1991) and that the higher the degree of affectedness born by an argument NP, the more semantically prominent it becomes. Thus, Beavers' work provides the theoretical underpinnings

to relate the present investigation to recent developments in sentence processing, such as the role of argument prototypicality and prominence.

One important issue to point out in advance relates to the experimental paradigm applied in the following investigations: while the first part of this thesis revolves around the impact of affectedness on aspects of sentence processing as a semantic feature induced by *verbs*, in the experiments to follow I will actually make use of *nouns* to address the above questions. On a first glance, this move will certainly appear unorthodox, at the very least. However, the nouns I will exploit for the present purposes are German deverbal nominalisations ending in the suffix *-ung*, such as *Bewunderung*/admiration or *Ermordung*/assassination; these are derived from verbs and – crucially – can inherit the base verb’s relevant semantic structure (Grimshaw, 1990). At the same time, however, the mechanisms and structural resources available for argument linking within NPs headed by such nominalisations are often much more restricted than within full sentences (Stiebels, 2006).

With German *-ung*-nominalisations of syntactically transitive verbs with a subject and a direct object, this leads to an interesting conflict when it comes to the licensing of either of these arguments as a genitive NP following the nominalisation, such as *des Präsidenten in die Bewunderung/Ermordung des Präsidenten* (*the admiration/assassination of the president*): whereas with *Bewunderung* the genitive NP *des Präsidenten* can potentially be understood as the event participant corresponding to the subject of the base verb *bewundern* (i.e. the ‘admirer’) *or* the one corresponding to the direct object (i.e. the one being admired), it can *only* be interpreted as the participant realised as the object with *Ermordung* (i.e. as the one being assassinated). Critically, the regulation of this linking pattern has been argued to depend upon whether the argument realised as object of the base verb counts as affected (as with *Ermordung*) or not (as with *Bewunderung*) (Ehrich & Rapp, 2000). Further, the affectedness status of the object argument provides the only relevant cue for resolving this argument realisation conflict, since morphosyntactic cues like case or number agreement are neutralised. Thus, somewhat paradoxically, NPs headed by *-ung*-nominalisations provide a context where the impact of the verb-induced semantic feature of affectedness for the regulation of argument linking processes is maximised – in the best case, focussing on such contexts in the experiments may thus represent a shortcut to addressing the questions raised above, rather than a detour. However, this approach also comes with risks and raises the question of the validity and generalisability of the results to other linguistic contexts (see the discussions in section 2.6 and chapter 5).

The structure of the first part of the thesis is as follows: chapter 1 introduces the theoretical background, including relevant aspects of argument realisation, the semantic structure of verbs and affectedness in particular. Concerning the latter, I will focus on the graded approach to affectedness of Beavers (2010, 2011) and how it opens up a perspective upon affectedness as a type of semantic prominence, and I will discuss correlating phenomena such as semantic transitivity and lexical aspect. Chapter 2 then reviews theoretical issues related to the impact of verbal semantic structure in deverbal nominalisations, especially on the role of affectedness for argument linking processes in German *-ung*-nominalisations. It further presents a first acceptability judgement study which is not only intended to provide initial

behavioural evidence, but also to validate the experimental paradigm. Note that these first two chapters are in part rather technical; section 2.5 presents an excursus on the syntactic status of genitive arguments and may be skipped safely by linguistically less inclined readers without running the risk of missing information crucial for the understanding of subsequent parts. Chapter 3 reviews the relevant psycholinguistic background for the self-paced reading and EEG experiments presented and discussed in chapter 4. Chapter 5 discusses the experimental findings in a wider context and concludes the first part.

## Part 2

The motivation for developing the software packages presented in the second part stemmed from the attempt to provide a maximal degree of methodological consistency throughout the different experiments, with the original challenge consisting in accounting for the effects of continuous, numeric variables: for largely practical reasons rooted in the experimental approach taken, I reoperationalised affectedness from Beavers' view as a graded hierarchy with four discrete levels to a continuous variable I will refer to as 'affectedness index' (AI), which is defined on a numeric scale. While this reoperationalisation of the main variable of interest simplified the design of the experiments, the standard analysis of variance (ANOVA) approach to the statistical analysis of behavioural as well as event-related potentials data (ERPs, which are derived from an EEG data set) cannot handle continuous numeric variables in the best way possible. Rather, statistically accounting for effects of such variables is most naturally done with regression-based tools, such as linear mixed effects models (LMMs), which have gained considerable ground in the analysis of behavioural data in the area of psycholinguistics (and cognitive psychology) in recent years (see Baayen, Davidson, & Bates, 2008, for example).

In addition to being able to handle variables with discrete levels as well as continuous variables (and any interactions thereof), LMMs offer the possibility to account for different sources of random variability, such subjects or experimental items, in one and the same model. Further, LMMs do not require prior averaging within subjects or items, thus making use of the full information present at the single-trial level of a data set. Thus, LMMs are the method of choice for the analysis of the behavioural reading time and acceptability judgement data in the present work.

In contrast to the behavioural data types, ERPs represents multidimensional data sets which are recorded at hundreds of sampling points in time and dozens (or even hundreds) of electrodes in space. Such spatiotemporal data sets thus hold extremely rich information, with the number of single trials in a typical psycholinguistic study ranging into the thousands. Within the standard ANOVA approach, a considerable amount of this information is discarded by not only averaging within subject and/or items prior to analysis, but also by extracting features (such as means or peak values) within predefined temporal windows and spatial regions of interest (ROIs). In addition, the same issues concerning the treatment of continuous variables within the ANOVA framework encountered with behavioural data arise in the context of ERP

data analysis. Yet, ERP studies compensating for some of these shortcomings by using LMMs are still relatively rare (see section 6.4.2), which may partially be due to the lack of software offering user-friendly ways of applying LMMs specifically for the analysis of ERP data.

The two software packages for the open statistical environment R introduced in part 2 are intended to address this gap by providing user-friendly interface software for the WFMM of Morris and Carroll (2006), which represents an extension of LMMs for *functional* – i.e. curve – data in the Bayesian framework (see Kruschke, 2010a, for an overview of the Bayesian approach to data analysis). One of the key advantages of the WFMM is its ability to model the effects of multiple categorical and/or continuous predictors along the whole extent of the curves (which can also be multidimensional curves, such as 2d-images) without extracting features such as means or peaks within predefined areas, at the same time offering the advantages of LMMs mentioned above. While this tool is available as free software, it has been lacking a user-friendly high-level interface, thus lowering its accessibility for potential users.

The *wrapfmm* package for the R language provides a wrapper for the core WFMM software which provides such a user-friendly interface and allows application of the WFMM for the analysis of general curve data. While it does not yet provide access to all of the functionality offered by the core WFMM software, it also extends and complements its functional scope; among the functions offered are functions for plotting and exploring a WFMM's underlying curve data as well as its posteriors in very flexible ways.

The *stepmom* package (**s**patiotemporal **e**lectrophysiological **m**odel **m**aps) builds upon *wrapfmm* and offers parallel functions adapted for ERP data. At its core are functions for using the WFMM to model ERP data without defining any temporal windows or spatial ROIs and extracting features within these; as a result, effects of predictors can be traced in the natural spatiotemporal domain of the ERP data at the original resolution. This approach considerably augments the amount of information for estimation as well as for the transmission of the results of an ERP experiment for the readership. As I will discuss in part 2, the packages as presented here also have their limits, which are largely imposed by computational issues and by the fact that the *wfmm* method in the *stepmom* package makes use of independent WFMMs for each electrode, thus making use of curve-internal correlation structure on the temporal axis for estimation, but failing to account for correlational structure on the spatial axis. Nevertheless, *stepmom* offers interesting functionality for exploring one's ERP data as well as modelling results flexibly with a number of methods which allow to plot many quantities of interest as (static or animated) topographic voltage maps or curve plots. In addition, I will present a novel type of posterior probability used for inference; these '**multiplicity-induced band of no difference scores**' (MULTIBONDS) are a natural by-product of a model's posteriors and reduce the risk of false positives by taking multiplicity along the extent of the curves into account (see section 7.7.2).

Thus, for my own ERP data, use of the *stepmom* package allowed me to apply the same design in all experiments, accounting for the continuous character of the AI variable in a natural way in all cases. Further, this modelling approach complements the partially exploratory character of the ERP experiment well.

Chapter 6 begins the second part and presents background information about relevant issues, including the advance of LMMs as an alternative to the conventional ANOVA approach to data analysis, current standard approaches to curve- and ERP-data analysis and of some more recent alternatives and the architecture of the WFMM of Morris and Carroll. In this chapter, I also intro the notion of 'holism' as a criterion for assessing the ability of a model to account for the full amount of structure present in the data at multiple levels, including that inherent to different functional data types, such as structure present at the temporal or spatial axes. Chapter 7 then introduces the *wrapfmm* package, providing detailed information about its current scope and limitations as well as a tour of the most relevant functions; in addition, it presents information about the novel posterior probability type, the MULTIBONDS. The following chapter 8 presents the *stepmom* package for ERP data analysis, again giving an overview of the most important functions. Chapter 9 discusses a number of issues related to the two packages, including current practical limitations and challenges of the presented versions in meeting the criterion of 'holism'.



## **Part I**

# **Affectedness as a factor in online argument linking: experiments in the nominal domain**



## Chapter 1

# Affectedness as a pivot between predicate meaning and syntactic structure

In this first chapter, I will introduce a number of relevant aspects of argument linking from the point of view of theoretical linguistics, setting up the theoretical context for the experimental investigations presented in the following chapters. After an introduction of basic aspects of and approaches to argument linking, I will focus on affectedness as a type of predicate-induced semantic argument prominence. Section 1.1 provides an introduction to some basic questions related to argument linking and to the role of affectedness as a semantic determinant in the mapping from verbal meaning to syntactic structure and section 1.2 sets up the general distinction between grammatically relevant and irrelevant aspects of verb meaning. Section 1.3 reviews the traditional approach to argument linking, which makes use of thematic roles and thematic hierarchies, and discusses its shortcomings. In section 1.4, I introduce alternative accounts of argument linking, concentrating on predicate decompositions as well as the Proto-role approach of Dowty (1991) and some related proposals. Finally, section 1.5 shifts the focus on affectedness as a graded phenomenon; it outlines the recent scalar approach to affectedness and its role in argument linking as presented by Beavers (2010, 2011), how this view allows to treat affectedness as a form of semantic prominence and how it relates to the approaches to argument linking discussed in the previous sections. In addition, it discusses the relation of affectedness with transitivity and lexical aspect, as these phenomena are tightly correlated in many aspects.

## 1.1 Argument linking: some basics

To illustrate some basic concepts and issues related to mapping from the meaning of a verb to syntactic structure, let us expand upon with the basis provided by examples (1)–(3) given in the preceding Introduction. The systematic contrast illustrated by the transitive (a) and the intransitive (b) variants of (1) and (3) is known as the *causative alternation*, since the meaning of the transitive versions can roughly be described as ‘X causes Y to V-intransitive’, where *X* refers to the participant encoded as the subject and *Y* to the one expressed as the direct object. The observation that such syntactic differences may be traced back to a semantic feature inherent to change-of-state verbs but absent in surface-contact verbs is strengthened by further patterns identified by Fillmore (1970, 1977):

- (6) a. I hit/slapped/struck his leg. *surface-contact verbs*  
b. I hit/slapped/struck him on his leg.  
c. I broke/bent/shattered his leg. *change-of-state verbs*  
d. \* I broke/bent/shattered him on his leg.
- (7) a. Mary hit the fence with the stick. *surface-contact verbs*  
b. Mary hit the stick against the fence. (= 7a)  
c. Mary broke the fence with the stick. *change-of-state verbs*  
d. Mary broke the stick against the fence. (≠ 7c)

In example (6), we find an alternation which is licensed for surface-contact verbs, but not for change-of-state verbs: when the direct object of a surface-contact verb is a body part (6a), the sentence can be paraphrased with the body part embedded within a prepositional phrase as in (6b). Such a paraphrase results in unacceptability for change-of-state verbs, however, as witnessed by examples (6c) and (6d). The examples in (7) are all grammatical, but while (7a) and (7b) are semantically equivalent, (7c) and (7d) are not paraphrases of one another: in (7c) *the fence* is wrecked, while it is *the stick* in (7d).

Thus, there are two classes of verbs which are defined by the presence or absence of certain aspect of their meaning: a definite change of state implied for one of their arguments. Further, a look at the German equivalents of (1) and (2) suggests that these semantically motivated regularities in the syntactic expression of verb arguments are not restricted to English. Speakers of German may talk about the same event using the following two sentences, which are the equivalents of sentences (2a) and (2b), marking the subject *der Junge* with nominative case, the direct object *das Fenster* with accusative and embedding *einem Stein* within a prepositional phrase:

- (8) a. Der Junge zerbrach das Fenster mit ein-em Stein.  
the.NOM boy.NOM broke the.ACC window.ACC with a-DAT rock.DAT  
‘The boy broke the window with a rock.’

- b. Der Junge traf das Fenster mit ein-em Stein.  
 the.NOM boy.NOM hit the.ACC window.ACC with a-DAT rock.DAT  
 'The boy hit the window with a rock.'

The German counterparts of (2) suggest parallels between the syntactic environments licensed by equivalent verbs across languages: in German, we may also express *das Fenster* (*the window*) as the sole participant of the *zerbrechen/breaking* event by marking it with nominative case (9a). However, doing the same with the verb *treffen* (*hit*) again leads to unacceptability (9b), just as in the English counterpart (2b).

- (9) a. Das Fenster zerbrach.  
 the.NOM window.NOM broke  
 'The window broke.'  
 b. \* Das Fenster traf.  
 the.NOM window.NOM hit  
 'The window hit.'

The linguistic study of systematic regularities behind the way in which arguments of verbs can be expressed on the clausal level is known as *argument realisation*, *argument mapping* or *argument linking*. Levin and Rappaport Hovav (2005, p. 3) take 'argument realisation' to encompass the complete range of phenomena related to the projection of verbal arguments to the syntax, while they understand 'argument linking' to be a concept somewhat more restricted in scope. While terminological details should not matter too much for the current purposes, I will mainly use the term (argument) linking in the remainder of this thesis, since it matches up with its use in psycholinguistics and since, in fact, the current focus is a very narrow one including only the core aspects of argument realisation.

At the heart of investigations into argument linking lies a presumably simple question: what are the principles that match up a verb's argument(s) with grammatical functions such as subject, direct and indirect object by morphosyntactic means? A syntactic strategy for grammatical function marking is found in English, where subjects are identified by their preverbal position, while objects are linked to a postverbal position. Via these positional cues, we can safely determine the subject and the direct object in simple active transitive sentences such as (10a). In German, on the other hand, the decisive cue to the grammatical status of a verb's arguments within a sentence is provided by morphological case marking: subjects are marked with nominative and direct objects, for example, with accusative case, as in (10b).

- (10) a. The policeman killed/captured/observed/saw the gangster.  
 b. Der Polizist tötete/fasste/überwachte/sah den Gangster.  
 the.NOM policeman.NOM killed/captured/observed/saw the.ACC gangster.ACC  
 'The policeman killed/captured/observed/saw the gangster.'

Via these morphosyntactic cues and general principles of argument linking, the semantic role of each of the arguments can be recovered. Let's assume that one of these principles

says something like: 'for verbs with two arguments, realise the *Agent* argument – the participant which consciously controls and executes the action denoted by the verb – as the (preverbal/nominative bearing) subject; realise the *Patient* argument – the participant which suffers the effect of the action and which undergoes a change of state – as the (postverbal/accusative bearing) direct object.'

For verbs such as *kill/töten* and *capture/fassen* in (10a) and (10b), such a rule should work well, since we can nicely attribute the Agent/Patient arguments to *the policeman* and *the gangster*, respectively. For *observe/überwachen* and *see/sehen*, however, the clear-cut dichotomy between Agents and Patients begins to blur; for a verb such as *observe/überwachen*, it is not clear at all whether the argument expressed as the direct object (*the gangster/den Gangster*) suffers any effect or undergoes some change of state due to the *observing* event. Attributing such a (loosely defined) Patient role to *the gangster/den Gangster* is at least equally problematic with a verb like *see/sehen*, where we are additionally faced with the difficulty of actually verifying that the argument linked as the subject (*the policeman/der Polizist*) is actually (something like) an Agent. After all, it is debatable whether a participant who sees somebody/something else has some kind of conscious control over the seeing event and whether there is anything to 'execute' at all with a verb like *see*.

This brief discussion brings us right to some of the most basic questions in the field of argument linking, such as: for transitive verbs which imply a clear Agent/Patient distinction between their two arguments, HOW COME the Agent is associated with the grammatical function of subject and the Patient with that of direct object in the first place? Whatever these basic principles are, how do they mediate argument linking in cases where arguments do not fit the Agent/Patient roles as well? How useful are distinct, categorical notions such as 'Agent' and 'Patient' for mapping from verbal lexical semantics to syntax to begin with? If such roles are necessary elements of a theory of argument linking, how are they best to be defined? Note that these issues are also of relevance to the question of how children exploit semantic and structural aspects of verbs during language acquisition (see Gleitman, 1990; Gropen, Pinker, Hollander, & Goldberg, 1991; Pinker, 1994).

In fact, investigating such questions about central principles of argument linking opens up a window onto the interface between lexical semantics and core syntactic structure – semantically, a verb and its arguments constitute the 'nucleus' of a sentence by providing the basic features of the event encoded as well as the participants involved in it. Syntactically, these act as the most basic constituents of a clause.

In a way, these questions are thus still more elementary than the more subtle puzzles related to argument alternations such as illustrated by the examples at the beginning of this section. Nevertheless, both aspects of argument realisation can be traced back to the same basic principles and aspects of a verb's meaning. Thus, the question whether one of the event participants undergoes some well defined change of state, for example, does not only matter for licensing certain kinds of argument alternations, but also figures into aspects of core argument linking.

In some theories of argument realisation, the concept of 'change of state' has been cast in terms of *affectedness*. Anderson (1979) made use of the notion of affectedness to explain

syntactic processes (NP preposing) within noun phrases: in example (11a), the NP *the city* counts as affected since it acts as complement of the head noun *destruction*, so it can occur within a *by*-prepositional phrase (PP) following *destruction* as well as in the form of a genitive in a 'preposed' position preceding *destruction*. In (11b), on the other hand, occurring with the predicate *knowledge, history* is unaffected and can thus not occur as a genitive preceding *knowledge* (cf. Anderson, 2006, pp. 121–122).

- (11) a. the destruction of the city/the city's destruction  
 b. the knowledge of history/\*history's knowledge

Other researchers have followed Anderson in exploiting the concept of affectedness or derivatives thereof for explaining a number of grammatical phenomena around nouns as well as verbs from the point of view of linguistic theory (e.g. Beavers, 2010, 2011; Jaeggli, 1986; Næss, 2007; Rozwadowska, 1988; Tenny, 1992, 1994) or language acquisition (Gropen et al., 1991, for example). Given the notion of change of state, the idea behind the semantic concept of affectedness is conceptually easy to grasp, but notoriously hard to define exactly without invoking purely intuitive notions, as noted by Beavers (2011). For the time being, I will stick to such a rather loose definition and refer to affectedness as implying some kind of change of state in or impingement upon one event participant; section 1.5 below will have a more detailed look at this concept.

The brief discussion of argument linking and alternations above suggests that affectedness plays a prominent role at the interface of lexical verbal semantics and syntactic structure; in fact, it is one of a small number of semantic features of predicates which enjoy a special status at the syntax/semantics interface by being 'grammatically relevant'. Before I discuss different approaches to linking and affectedness in more detail, the next section will explain what it takes for a certain aspect of a word's meaning to count as grammatically relevant.

## 1.2 Grammatically relevant aspects of verb meaning

As Grimshaw (2005b, p. 75) aptly notes, "[i]t is very easy to ask what a word means, and very hard to answer the question". Take the nouns *dog*, *horse*, *duck* and *snake* as an example and suppose a naive speaker of English is asked to explain the most conspicuous features of the meaning of each one and how they differ from each other. The hypothetical speaker may rightfully point out that all refer to animals, that *dog* and *horse* refer to creatures with four legs, but *ducks* have only two and *snakes* none and that *dogs* bark, *horses* neigh, *ducks* quack and *snakes* hiss; she may add that only *ducks* can fly and all *dogs* as well as *horses* have fur, while *ducks* have feathers and *snakes* neither; that *horses* are bigger than all the other animals in the list, that some people consider *dogs* (or maybe *horses*) their best friends and that some of the four animals are more likely to be eaten by humans in her particular culture. There are many more features that come to mind and the list is potentially open-ended. And while all of these characteristics matter for our intuitive definition of the meaning of these words, it may be surprising for our naive speaker to learn that none of them

make actually a difference within the *linguistic system* – in fact, from the point of view of the grammatical system, all of these words are in a way indistinguishable.

This discrepancy between what we, as naive speakers of a language, intuitively find important and distinctive about word meanings and what (most) linguists consider to be relevant semantic aspects of a word stems from the finding that a lot is to be gained for the study of the lexicon/syntax interface by separating two components of word meaning: those semantic features which actually have repercussions in syntax – for verbs especially those aspects which influence argument realisation options – versus those which don't (e.g. Grimshaw, 2005b; Jackendoff, 1990; Levin & Rappaport Hovav, 2005; Pinker, 1989; Rappaport Hovav & Levin, 1998; Talmy, 1985; Van Valin Jr., 1990, inter alia). The former aspect is known as 'grammatically relevant' (Levin & Rappaport Hovav, 2005; Pinker, 1989), 'structural' (Rappaport Hovav & Levin, 1998) or 'semantic structure' (Grimshaw, 2005b) and the latter as 'idiosyncratic' (Rappaport Hovav & Levin, 1998) or 'semantic content' (Grimshaw, 2005b).

Grimshaw (2005b, p. 75) provides examples illustrating the difference between what she refers to as the 'semantic structure' (grammatically relevant) and 'semantic content' (grammatically irrelevant) of a verb:

- (12) a. He melted the ice-cream.  
b. He froze the ice-cream.

- (13) a. She wrote a book.  
b. She drew a picture.

Grimshaw (2005b, p. 76) refers to these examples in conveying the gist of the partition between semantic structure and content (emphasis in the original):

"To preview the conclusion, I will split the analysis of a verb like *write* in the following way: that *write* means to do something and not to be something is linguistic; that it means what it means and not what *draw* means, is not. Similarly, that *melt* and *freeze* both (can) mean to change state is linguistic, that they concern changes in liquidity, and that each means what it means and not what the other means is not: it is purely a matter of content. The aspect of meaning that distinguishes *write* from *draw*, or *melt* from *freeze* is of no linguistic significance and plays no role in the grammatical system of a language."

Some more examples will help to further illustrate the difference between these two components of meaning for verbs. The following examples are taken from Levin and Rappaport Hovav (2005, examples (1)-(4) on page 11) and all involve verbs of sound emission:

- (14) a. The truck rumbled.  
b. \* Peter rumbled the truck.

- (15) a. The tea kettle whistled.



- b. \* The boiling water whistled the tea kettle.
- (16) a. The tea cups clattered.
- b. I clattered the tea cups as I loaded the dishwasher.
- (17) a. The windows rattled.
- b. The storm rattled the windows.

In their intransitive use, all four verbs of sound emission appear with the sound emitter realised as the subject (the (a) examples), but only *clatter* and *rattle* allow transitive, causative variants (as in examples (16b) and (17b)) in which the sound emitter is linked as the direct object and the argument in subject position refers to a natural force or entity causing the sound emission. As pointed out by Levin and Rappaport Hovav, the ability of any of these four verbs to appear in a transitive causative construction does not depend on any of the respective resulting sound's inherent 'physical' real world properties, such as its volume, pitch or duration. Rather, what seems to make the difference is whether the sounds are produced internally to the sound emitter (such as *rumble* and *whistle*) or externally (*clatter* and *rattle*): only if the sound is construed as being emitted externally to the emitter, the respective verb lexicalises an externally caused event which can potentially be controlled by some external force and may thus appear in a causative construction (Levin & Song, 1997).

This brief case study also illustrates how linguists go about isolating the set of grammatically relevant components of verb meaning: verbs which group together grammatically by exhibiting the same argument realisation options should have a critical semantic feature in common. If such a shared feature is identified for a group of verbs showing common argument realisation patterns, the respective verbs may of course further cross-classify into other groups with systematic and semantically motivated grammatical behaviour. The exact composition of the set of grammatically relevant features and their definitions may of course differ between different linguistic frameworks and approaches to argument realisation, but some are more persevering than others and have been implicated again and again in a range of phenomena related to argument mapping. Among these is affectedness (e.g. Anderson, 1979, 2006; Beavers, 2010, 2011; Gropen et al., 1991; Jaeggli, 1986; Næss, 2007; Rozwadowska, 1988), which I will take a closer look at in section 1.5 below.

So far, I have focussed on the partition between semantic structure and semantic content (in Grimshaw's sense) from the perspective of lexical semantics and argument mapping. However, it also matters at the level of combinatorial semantics, when words are combined into phrases and sentences and the complex meaning of the whole expressions are computed from the meanings of the single lexical items within the expression (Grimshaw, 2005b; Pytkäinen, Brennan, & Bemis, 2011). Pytkäinen et al. (2011) draw attention to the fact that much work done within the cognitive neuroscience of language has actually focussed on the role of semantic *content* in single word and sentence processing, rather than on the mechanisms behind what most semanticists consider central processes of combinatorial semantics. This difference in focus has not only led to terminological confusions, but also to misconceptions about the functional role of prominent indexes of 'semantic' processing such as the

N400 component known from the electrophysiological literature on word and sentence processing (see also section 3.2.3 for more discussion).

To illustrate, the following examples are from Grimshaw (2005b, pp. 79–80); the sentences in (18) violate what is known as ‘selectional restrictions’ and appear (at least) weird, since we cannot drink non-drinkable things such as cars, nor slice liquids. However, Grimshaw’s point is that selectional restrictions fall into two classes, one based on semantic content, the other on semantic structure. The oddness of sentences as in (18) stems from violations of restrictions induced by the semantic content of the verbs – from the point of view of linguistically defined semantics, however, nothing is wrong about these sentences, since all of the object NPs are actually perfectly free to combine with a verb like *drink* or *slice*. This is made clear by the examples in (19), where the highly unlikely character of the events expressed in (18) is asserted, which renders these statements perfectly normal.

- (18) a. # He drank the meat/car/universe.  
b. # He sliced the orange juice.
- (19) a. Noone can drink meat/cars/the universe.  
b. It isn’t possible to slice orange juice.

Thus, the strangeness of the sentences in (18) is not a linguistic issue, but it comes from a clash of *conceptual* knowledge about what it ‘means’ to drink or slice something and about what kinds of things are drinkable and sliceable – in the words of Pykkänen et al. (2011, p. 1323), such sentences are “not semantically ill-formed, but rather just ill-fitting to our world knowledge”. To keep these two aspects of word meaning apart, I will use ‘semantic’ in referring to the linguistically relevant aspects of meaning (i.e. semantic structure) and ‘conceptual’ to those aspects which are inert within the linguistic system (i.e. semantic content).

On a final note, the distinction between grammatically relevant and irrelevant aspects of meaning also matters for word classes other than verbs. Going back to the initial examples *dog*, *horse*, *duck* and *snake*, they are probably all alike for the grammatical system, since they are identical in terms of semantic structure. Nouns like *cows* and *cattle*, on the other hand, are largely indistinguishable in terms of semantic content (i.e. conceptually), but differ in semantic structure, since the former is the plural form of a count noun and the latter is a mass noun, a semantic difference which is accessed by the linguistic system. This point is made by Grimshaw (2005b) together with the observation that verbal semantic structure can also occur in nominals derived from verbs; deverbal nominalisations such as *destruction* (derived from the verb *destroy*) can thus inherit a verb’s semantic structure, a point which is crucial for the paradigm applied within this thesis and further discussed in the following chapter 2 (see also Grimshaw, 1990).

## 1.3 Thematic roles and argument linking

### 1.3.1 Thematic roles

Distilling grammatically relevant aspects of word meaning is one step in uncovering principles of argument realisation; another important step is to somehow structure and organise these features in a way which allows to develop an efficient syntax/semantics interface which parsimoniously captures crucial mechanisms of the mapping from semantic arguments of a verb to syntactic functions. One prominent approach involves *thematic roles*, which “provide one way of relating situations to their participants” (Davis, 2011, p. 400): thematic roles are taken to be *discrete and atomic* (i.e. semantically unanalysable) concepts/labels which are assigned by a verb to its arguments and define the role a given argument plays in the situation/event denoted by the verb (e.g. Fillmore, 1968, 1970; Jackendoff, 1972) – for overviews, see Davis (2011) and Levin and Rappaport Hovav (2005, chapters 1,2 & 6).<sup>1</sup> Typically, a verb’s lexical entry will list one thematic role for each of its arguments; such lists are termed *semantic role lists* by Levin and Rappaport Hovav (2005) and are also known as ‘case frames’ Fillmore (1968) or ‘theta-grids’ (Stowell, 1981).

Crucially, the set of thematic roles is usually assumed to be fairly small in size – while the exact inventory of thematic roles as well as their definitions have been a matter of intense debate (without any consensus, as pointed out below), in the following example (20), I list a number of frequently encountered roles together with rough characterisations.<sup>2</sup>

- (20) a. *Agent*: initiator of action, capable of volition; affecting another participant  
b. *Patient*: participant being affected/undergoing some change of state  
c. *Theme*: participant undergoing some caused change of location/motion  
d. *Experiencer*: participant experiencing some mental effect/undergoing a mental change of state  
e. *Instrument*: intermediary/means (put to use) in performing an action  
f. *Source*: the place from which somebody/something moves  
g. *Goal*: the place to which somebody/something moves  
h. *Location*: place of object or action

The promise that such thematic roles hold for structuring the syntax/semantics interface is to package semantic similarities of verb meanings reflected in their arguments in a way

---

<sup>1</sup>As Manfred Krifka has pointed out to me, the basic concept of thematic roles can be traced back as far as Pāṇini’s grammar of Sanskrit (ca. 5<sup>th</sup> century BC), which classified participants of an action in terms of six distinct *kāraṅkas*. Kiparsky (2009) lists these as Agent (*kartṛ*), Goal (*karman*), Recipient (*saṃpradāna*), Instrument (*karana*), Locative (*adhikaraṇa*), and Source (*apādāna*). In functional terms, Kiparsky notes that “the *kāraṅkas* mediate between meaning and morphosyntactic surface structure” (p. 50; emphasis in the original), thus coming quite close to the modern notion of thematic roles.

<sup>2</sup>The definitions are intended to be roughly representative of frequently found features characterising a given role and mainly serve illustrative purposes – they are NOT intended to represent the inventory and definitions of a specific approach, however. Some of the definitions are based on those listed by Palmer, Gildea, and Xue (2010, table 1.1 on page 4).

which allows to group verbs into different natural classes whose members systematically share options for argument realisation. To illustrate how thematic roles can be brought to bear upon argument linking patterns, consider again the verbs *break* and *hit*. While deviating somewhat in terminology, we may follow Fillmore (1970, p. 131) and assume that these verbs' lexical entries feature the following role lists:

- (21) a. *break*: Agent, Patient, Instrument  
 b. *hit*: Agent, Place, Instrument

Note that the two lists differ in listing a Patient and a Place role for their second argument, respectively, i.e. the two verbs differ semantically. The next two examples are in fact (1) and (2) repeated here. The sentence pair in (2) shows again that the two verbs also differ syntactically in their argument realisation options. The differences in argument realisation patterns can now be derived by subject/object selection rules along the following lines (cf. Fillmore, 1970, pp. 128–129): for *break*, the Patient argument must be expressed, while the Agent and Instrument arguments may be realised. If the Agent is expressed syntactically, it is linked as the subject, as in (1a). If the Instrument argument is expressed, it is realised as the subject ONLY if the Agent is NOT expressed within the sentence (example (22a)) – if the Agent is realised, however, the Instrument is linked obliquely within a PP (as in example (1a)). If NEITHER the Agent nor the Instrument are realised, the Patient becomes the subject, as in (2a). The Patient is linked as the direct object if either the Agent or Instrument are expressed syntactically (see (1a) and (22a), respectively). For *hit*, on the other hand, we may assume that the ungrammaticality of (2b) stems from the difference in the composition of the semantic role lists, which leads to the obligatory expression of either the Agent or Instrument arguments, rendering the sole realisation of the Place argument as the subject illicit.

- (1) a. The boy broke the window with a rock.  
 b. The boy hit the window with a rock.
- (2) a. The window broke.  
 b. \* The window hit.
- (22) a. The rock broke the window.  
 b. The rock hit the window.

Crucially, the semantic role lists given in (21) as well as the above rules can now be extended to the other verbs patterning together with *break* (e.g. *bent*, *folded*, *shattered*, *cracked* – (3)) and *hit* (such as *slapped*, *struck*, *bumped*, *stroked* – (4)), respectively. Thus, the correlations between lexical semantics of groups of verbs and their argument linking patterns can be systematised via thematic role lists and linking rules.

To arrive at a predictive and parsimonious theory of argument realisation using thematic roles, it is generally agreed that the set of roles must be limited in size, which – in turn – implies

that each role must allow a large number of related arguments to fit its definition. Thus, finding the right 'grain size' in choosing the roles and their definitions is critical for any model of argument linking making reference to thematic roles – however, as discussed by several authors (e.g. Ackerman & Moore, 2001; Dowty, 1991; Levin & Rappaport Hovav, 2005), no consensus on these aspects has emerged yet.

While the notion of 'being an Agent' is relatively intuitive, for instance, defining an Agent's properties has proven surprisingly difficult. As a case in point, Dowty (1991), Ackerman and Moore (2001) and Levin and Rappaport Hovav (2005) all cite Cruse (1973), who splits 'the' Agent role into the four 'Agent-like' subroles *Volitive*, *Effective*, *Initiative* and *Agentive*. According to this four-way split, since *John* in (23a) is wilfully involved in the event of tearing, he counts as a Volitive, while it would bear the Agent role following more standard definitions; the *machine* in (23b), on the other hand, counts as Agentive according to the definition of Cruse, but may qualify as Instrument according to the rough role characterisations provided in (20).

- (23) a. John tore the document to pieces.  
b. The machine shredded the document to pieces.

The lack of widely accepted role inventories and definitions has led to a rather loose application of role labels in many cases and some of them are often used as defaults for arguments whose role eludes straightforward classification – as Jackendoff (1987, p. 381) remarks, "[s]ome have taken Theme or Patient to be such a default role: if one can't think of anything else to call an NP, call it Theme or Patient (and some have treated the two terms as interchangeable)". Also, as pointed out by Levin and Rappaport Hovav (2005, pp. 27–28), there is no easily definable relation between the syntactic function of direct object and its semantic complement in terms of thematic roles. While 'true' Patients (and/or Themes) are cross-linguistically linked as direct objects with syntactically transitive verbs, a whole number of other roles may be realised as direct objects in many languages. Levin and Rappaport Hovav (2005, p. 28) list the examples in (24) (the references are provided in the original).

- |  |   |
|--|---|
| (24) a. The engineer cracked the bridge. | <i>Patient</i>                                    |
| b. The engineer destroyed the bridge.    | <i>Patient/consumed object</i>                    |
| c. The engineer painted the bridge.      | <i>incremental theme; Dowty (1991)</i>            |
| d. The engineer moved the bridge.        | <i>theme</i>                                      |
| e. The engineer built the bridge.        | <i>effected object/factitive; Fillmore (1968)</i> |
| f. The engineer washed the bridge.       | <i>location/surface</i>                           |
| g. The engineer hit the bridge.          | <i>location; Fillmore (1970)</i>                  |
| h. The engineer crossed the bridge.      | <i>path</i>                                       |
| i. The engineer reached the bridge.      | <i>goal</i>                                       |
| j. The engineer left the bridge.         | <i>source</i>                                     |
| k. The engineer saw the bridge.          | <i>stimulus/object of perception</i>              |

l. The engineer hated the bridge.

*stimulus/target or object of emotion*

Parallel issues arise with the use of the label 'Agent' for subjects and the range of semantic roles that may be linked as subjects; Ackerman and Moore (2001, p. 23) note that "rather than a label such as AGENT being empirically motivated as the result of employing [a principled] algorithm, it is applied *post-hoc* to class together entities that are known to share some important behaviors" (emphasis in the original). The following examples (Ackerman & Moore, 2001, p. 23) in (25) all have subject arguments, but "it is not evident that these arguments share any semantic property. Labeling them all AGENT is tantamount to simply stipulating that the subject of active clauses is equivalent to being an AGENT" (Ackerman & Moore, 2001, p. 23). One strategy might be to fragment thematic roles into subtypes along the lines of Cruse (1973), but if carried through, one runs the risk of ending up with what Dowty (1991) calls 'individual thematic role types', which amounts to assigning to the subject of a verb such as *break* a 'breaker' role and to that of *hit* the 'hitter' role. Obviously, however, such a fragmentation means to miss the chance of identifying those semantic commonalities present across arguments which allow to make generalisations about their linking behaviour.

- (25) a. The duck is swimming.  
b. The duck is dying.  
c. The duck swallowed the frog.  
d. The duck saw the frog.

Another problem with thematic roles discussed by Dowty (1991) arises with the roles assigned to argument of symmetric predicates, such as *resemble* (26) or *border* (27). In such cases, the two arguments appear to bear the same thematic role and it is not clear what (unique) event-based role to assign to each of the verbs' arguments in a way consistent with principles of thematic role-based linking principles.<sup>3</sup>

- (26) a. Nella resembles Luk. =  
b. Luk resembles Nella.  
(27) a. The field borders the wood. =  
b. The wood borders the field.

While thematic roles have been the predominant means of addressing issues related to argument realisation throughout the last decades, the above discussion points out some of the unresolved issues which have been plaguing this approach, the foremost of which being the question about the right grain size in choosing the set of relevant thematic roles and their definitions. Other problems inherent to thematic role-based theories of argument realisation are discussed by Dowty (1991), Levin and Rappaport Hovav (2005), Ackerman and Moore (2001), Davis and Koenig (2000) and Davis (2011), amongst others.

<sup>3</sup>Note that a similar issue arises with the verb pair *buying/selling*, since here "both buyer and seller must act Agentively (voluntarily) [...] and there is no obvious reason why either is entailed to act 'more Agentively' than the other" (Dowty, 1991, p. 556).

### 1.3.2 Thematic hierarchies and prominence scales

So far, I have implicitly presented thematic role entries as unstructured and unordered lists; however, if structure is imposed onto a semantic role list by arranging the respective thematic roles according to some kind of prominence relations, we can speak of thematic hierarchies. In (28), the Agent, Instrument, Patient and Theme roles are arranged in such a hierarchy, with '>' implying a higher position on the hierarchy (precedence) for the argument(s) to the left and '/' a tie between the respective arguments (the following examples (28)–(30) are based upon Levin & Rappaport Hovav, 2005, pp. 155–156).

(28) Agent > Instrument > Patient/Theme

The benefit to be gained from imposing hierarchical structure onto thematic role lists lies in the possibility to formulate linking rules in terms of the position of a role within the hierarchy, rather than by reference to specific role labels (e.g. Bresnan & Kanerva, 1989; Grimshaw, 1990). Note that the subject selection rule introduced above in the context of *break/hit* type verbs makes reference to concrete role labels such as Agent, Instrument and Patient and associates them with syntactic functions. In (29), I render a more compact version of the subject selection rule as presented by Fillmore (1968, p. 33), where 'A' stands for Agent, 'I' for Instrument and 'O' for object, which corresponds to the Patient (or Theme) role.

(29) If there is an A, it becomes the subject; otherwise, if there is an I, it becomes the subject; otherwise, the subject is the O.

Together with a hierarchically structured thematic list such as (28), this rule can now be derived by a simplified version of the above rule: the formulation in (30) does not refer to any specific thematic role labels, thus gaining in generalizability. The result is the same, since when an Agent is present, it will be realised as subject, while other roles can be linked as subject in the absence of Agents. Likewise, the priority of Instruments over Patients/Themes for subject realisation follows from the precedence relations made explicit by the hierarchy in (28).

(30) The argument of a verb bearing the highest-ranking thematic role is its subject.

Another plus of thematic hierarchies is the fact that they naturally handle what Levin and Rappaport Hovav (2005) refer to as 'context dependence': frequently, the way a specific argument is realised morphosyntactically is not determined on basis of the thematic role it bears, but rather depends on the roles borne by its co-arguments realised within a sentence – thus, an Instrument can only become subject if no Agent is realised concurrently (e.g. (31a)) and a Patient only qualifies for linking as subject if no other argument with a higher role (such as Instrument – (31b)) is realised (see Levin & Rappaport Hovav, 2005, section 6.2). In Fillmore's subject linking rule (29) above, context dependence is explicitly introduced by the if-else conditions, which make sure that higher roles take precedence for subject linking over lower ones. With a linking rule referring to a thematic hierarchy as in (30), context dependence is implicitly handled by the hierarchical ordering imposed onto the thematic roles.

- (31) a. \* The rock broke the window by John.  
 b. \* The window broke with a rock.

Further, thematic hierarchies provide a natural means to preserve the prominence relations between arguments inherent in the lexical semantics representation in the mapping to syntax; such a principle of 'prominence preservation' (for an overview see Levin & Rappaport Hovav, 2005, section 5.2) is assumed in a number of approaches to argument mapping (e.g. Ackerman & Moore, 2001; Belletti & Rizzi, 1988; Larson, 1990; Wunderlich, 1997). Often, lexical semantic prominence relations are mapped onto morphosyntactic hierarchies, such as grammatical function hierarchies or morphological case hierarchies as illustrated by (32) and (33) (cf. Levin & Rappaport Hovav, 2005, p. 142), respectively.

(32) subject > object > indirect object > oblique

(33) nominative > accusative > dative > oblique cases

While thematic hierarchies are promising constructs in the quest for economical principles regulating the mapping from lexical semantics to syntax, they inherit the problems inherent to thematic role-based approaches discussed in the previous section, i.e. the question of which roles to include in the inventory of a thematic hierarchy and how to define them. In addition, however, with thematic hierarchies there is no consensus yet as to the ranking of the roles within the hierarchy (see Baker, 1996, for example); while it is largely agreed that Agents are located highest on the thematic hierarchy and Patients/Themes are to be found at its lower end, there has been a lot of variability concerning the relative placement of Patients/Themes and certain other roles such as Goal and Source, for example. Levin and Rappaport Hovav (2005, pp. 162–163) list a representative sample of 16 different formulations of 'the' thematic hierarchy and often the differences arise from diverging views about assumptions of the mapping algorithm, the underlying syntactic structure and processes and/or the phenomena which are the focus of interest of a particular study.

Finally, Levin and Rappaport Hovav (2005, section 6.4) discuss the relation of thematic hierarchies to other types of 'prominence scales' frequently used for the explanation of a wide range of different grammatical processes (see Lockwood & Macaulay, 2012, for an overview), including aspects of argument mapping. Similar to thematic hierarchies, these rank specific linguistic features according to their values, where the order is sometimes thought to reflect some kind of cognitive salience. The following 'animacy hierarchy' represents an example taken from Comrie (1989, p. 128) (presented by Lockwood & Macaulay, 2012, p. 431):

(34) 1st/2nd person pronouns > other human NPs > animal NPs > inanimate NPs

While thematic hierarchies are sometimes assumed to be a member of the family of 'natural prominence scales', Levin and Rappaport Hovav caution against confounding possible effects of prominence scales implied in argument realisation processes with those actually induced by



the lexical-semantic properties of verbs which are sometimes cast within thematic hierarchies. They point out that the attributes typically included in prominence scales – such as person, animacy, definiteness and number – represent *properties of the NP which fills an argument position* of a verb, but “they should not, strictly speaking, be considered defining properties of the semantic role” (Levin & Rappaport Hovav, 2005, p. 165). Levin and Rappaport Hovav (2005, p. 171) note that the values of the properties of a specific NP can be related to a number of different hierarchies/scales and that often there is a tendency for them to align in position across scales:

“A given NP in a sentence usually has various attributes and, thus, is associated with values drawn from several natural prominence scales. The values of the various attributes associated with a given NP tend to be chosen from the same ends of the relevant scales. So the Agent role is ranked high on the thematic hierarchy, and Agents tend to be human, a value high on the animacy hierarchy; the Agent role is also usually associated with subject, which is high on the grammatical relations hierarchy, and with nominative case, which is high on the morphological case hierarchy.”

Levin and Rappaport Hovav (2005, p. 165, my emphasis) further point out a consequence of this kind of correlation in ordering of NP-related feature values across the different hierarchies: “since the properties of the role fillers may align in particular ways with the associated semantic roles, they may also give rise to what *appear* to be thematic hierarchy effects”. Thus, sometimes argument linking phenomena related to NP-inherent features ordered along prominence scales give rise to *apparent* thematic hierarchy effects. It is important to keep this in mind and try and keep apart effects related to attributes inherent to specific NP-fillers of verbal arguments (drawn from natural prominence scales) and those which can be traced back to the lexical semantics proper of the predicate. This distinction will also be relevant for the discussion of the findings of the current study and interpreting them with respect to those of other sentence processing studies (see the discussion in chapter 5, for example).

## 1.4 Alternatives to thematic role-based approaches

### 1.4.1 Predicate decompositions

As one alternative to thematic role-based approaches to argument realisation, some researchers have shifted attention away from the roles borne by arguments and have focussed on decomposing and structuring the meaning of verbs themselves, resulting in representations of verb meanings known under different names, among the most prominent of which are ‘event structures’ and ‘predicate decompositions’ (e.g. Jackendoff, 1990; Levin, 1999; Pinker, 1989; Rappaport Hovav & Levin, 1998; Wunderlich, 1997, – for an overview see Levin and Rappaport Hovav 2011). Such structured representations of verb meanings are composed of primitive predicates which are thought to encode grammatically relevant semantic elements. Just as the set of thematic roles is assumed to be restricted, the set of

these basic semantic predicates is limited to a few; commonly found predicates include are *ACT/DO*, *CAUSE*, *BECOME/CHANGE*, *GO*, *BE* and *STAY*, though additional ones have been proposed by some researchers (for discussion, see Levin & Rappaport Hovav, 2011, section 4).

These primitive predicates take arguments which represent event participants and may be combined to yield representations of semantically complex verb classes. While the details of such decompositions and the predicates used may differ across the various theories, I will exemplify the general approach as presented by Rappaport Hovav and Levin (1998). The examples in (35) (from Rappaport Hovav & Levin, 1998, p. 108) represent five different predicate decompositions, which also happen to correspond to the four major event types of activities, states, achievements and accomplishments (Dowty, 1979; Vendler, 1957). Note that the decompositions corresponding to activities and states in (35a) and (35b) are structurally simple, while others are complex; the representation of accomplishments in (35d), for example, combines the simpler ‘templates’ (35a) and (35c) and links them via *CAUSE*.

- (35) a. [*x ACT*<sub><MANNER></sub>] *activity*  
       b. [*x <STATE>*] *state*  
       c. [*BECOME* [*x <STATE>*]] *achievement*  
       d. [[*x ACT*<sub><MANNER></sub>] *CAUSE* [*BECOME* [*y <STATE>*]]] *accomplishment*  
       e. [*x CAUSE* [*BECOME* [*y <STATE>*]]] *accomplishment*

Such predicate decompositions capitalise on the distinction between the grammatically relevant and irrelevant aspects of verb meaning, or, in Grimshaw’s (2005b) terms, semantic structure and semantic content. In work on predicate decomposition, the part of the meaning which is idiosyncratic to each verb is also known as the ‘constant’ or ‘root’; in the decompositions in (35), the root part is actually integrated into the structural templates, as signalled by the italicised material enclosed within angled brackets. Rappaport Hovav and Levin (1998, pp. 108–109) assume that each root belongs to one of a small and fixed set of ontological types (such as state, thing, place or manner) and the root’s ontological category type determines its association with a specific predicate decomposition (or ‘event structure template’). Thus ‘manner of motion’ verbs such as *jog* and *hobble* as well as sound emission verbs such as *whistle* are linked to the ‘manner’ template in (36a) and causative verbs implying an externally caused change of state (such as *break* or *open*) with the decomposition in (36a) (examples are from Rappaport Hovav & Levin, 1998, p. 109).

- (36) a. manner → [[*x ACT*<sub><MANNER></sub>]  
           (e.g. *jog*, *run*, *creak*, *whistle*, ...)  
       b. externally caused state → [[*x ACT*] *CAUSE* [*BECOME* [*y <STATE>*]]]  
           (e.g. *break*, *dry*, *harden*, *melt*, *open*, ...)

Related verbs are thus grouped together into natural classes and verbs within a class can be distinguished via the idiosyncratic part contributed by each verb’s root. (37) lists the

decompositions for three distinct causative verbs (*break*, *open* and *shorten*): while all share the same structural template and contain a *BECOME* predicate which takes a state as its argument, the concrete nature of this state is determined by the contribution of each individual verb's root, i.e. the entry for *break* and *open* specify that the participant represented by the variable *y* ends up in a broken and open state, respectively.

- (37) a. *break*:  $[[x \text{ ACT}] \text{ CAUSE } [\text{BECOME } [y <\text{BROKEN}>]]]$   
 b. *open*:  $[[x \text{ ACT}] \text{ CAUSE } [\text{BECOME } [y <\text{OPEN}>]]]$   
 c. *shorten*:  $[[x \text{ ACT}] \text{ CAUSE } [\text{BECOME } [y <\text{SHORT}>]]]$

In these decompositions, a verb's arguments available for linking are encoded by the unfilled argument position opened up by the basic predicates, such as *x* and *y* in the above examples. While these arguments within the templates are unlabelled, it is possible to define thematic roles by virtue of being an argument of a specific basic predicate: Agents, for example, can be defined as being the first argument associated with a *CAUSE* predicate (i.e. the *x* in (37)) (Jackendoff, 1972), which means that thematic roles are viewed as derived rather than basic notions (Levin & Rappaport Hovav, 2005, p. 69).

A further advantage of predicate decompositions is that they have inherent hierarchical structure which can be referenced and used by rules and principles of argument realisation; rather than referring to concrete thematic roles, for instance, rules of linking can be formulated with reference to the 'geometry' of the structured semantic representations and the hierarchical position of the open participant arguments within it (e.g. Levin & Rappaport Hovav, 2011, section 6). The hierarchical nature of predicate decompositions also provides a natural means of accommodating prominence preservation in the mapping from lexical semantics to syntax and the role ranking provided by thematic hierarchies comes for free. As pointed out by Levin and Rappaport Hovav (2011, p. 433), "most work adopting a thematic hierarchy does not provide independent motivation for the posited role ranking. Predicate decompositions can provide some substance to the notion of a thematic hierarchy by correlating the position of a role in the hierarchy with the position of the argument bearing that role in a predicate decomposition."

The use of predicate decompositions in the explanation of argument realisation phenomena has proven an attractive alternative to approaches involving thematic roles and hierarchies. Rappaport Hovav and Levin (1998) show how basic core templates as presented in (35c) can be used to construct derived verb meanings, guided by a number of general principles. One of the applications which are of interest for the current study is presented by Levin (1999), who makes use of predicate decompositions to explain phenomena such as the obligatory vs. optional presence of object arguments and to define 'core transitive' verbs, a topic I will come back to in section 1.5.2.1 below.

## 1.4.2 Prototype-based argument linking approaches

### 1.4.2.1 Dowty's Proto-Roles

A number of researchers have pursued another alternative to argument realisation approaches based on atomic thematic roles by splitting up concrete roles into a number of features, thus letting go of the assumption that notions such as Agent and Patient are indivisible primitives. Rozwadowska (1988) suggests to decompose thematic roles into the features [+/- sentient], [+/- cause] and [+/- change]; in this framework, an Agent is defined as [+ sentient, + cause, - change] and a Patient bears the opposite feature specifications [- sentient, - cause, + change]. Other traditional thematic roles can be defined by different feature specifications. In a similar spirit, Næss (2007) proposes a system comprising the three features [+/- volitionality], [+/- instigation] and [+/- affectedness] and Reinhart (2003) makes use of a decomposition involving only two features [+/- c] ('cause change') and [+/- m] ('mental state'). While such feature decompositions go some way towards addressing the problem of the right grain of thematic roles (see Levin & Rappaport Hovav, 2005, section 2.3.1), other researchers have developed ways to employ feature decompositions without making recourse to thematic roles with sharp boundaries.

The best known of these approaches is Dowty's (1991) 'Proto-roles' theory: as with feature decompositions, Dowty defines a number of properties characterising prototypical Agents and Patients, which he views as lexical entailments imposed by a verb on its arguments. On this view, thematic roles are best understood as recurring clusters of these lexical entailments associated with certain groups of verbs and are thus emergent second order properties. In contrast to the feature decomposition approaches briefly mentioned above, however, Dowty argues that only two roles need to be recognised to explain core argument realisation patterns, Proto-Agent and Proto-Patient. As the names suggest, these roles are NOT intended to be defined by fixed and specific sets of feature values, but they are conceptualised as cluster-based, overarching *prototype* roles; a given argument of a verb might be a more or less prototypical member of the Proto-Agent role or a more or less representative member of the Proto-Patient category, in the spirit of Rosch (1973).

The properties which Dowty includes for defining either Proto-role are listed in table 1.1: while most of them are reminiscent of the properties encountered in other approaches to thematic roles, we have not encountered 'incremental theme' yet, which is based on aspectual notions and which will be discussed further in section 1.5.2.2 below.<sup>4</sup> An argument of a verb which either has all or most of the Proto-Agent properties listed in the Proto-Agent column of table 1.1 counts as central, prototypical instance of a Proto-Agent, while an argument exhibiting all/most of the Proto-Patient characteristics is most representative of the Proto-Patient cluster concept. Examples for prototypical Agents are *the extremist* in (38a) and *John* in (38b); *the president* in (38a) and *a house* in (38b) are instances of central members of the Proto-Patient concept.

---

<sup>4</sup>Note that Dowty parenthesises the last property within each column (does or does not exist independently of the event) since he is not sure whether these should be seen as being discourse-based characteristics, rather than event-based notions.

Proto-Agent	Proto-Patient
Volitional involvement in the event or state	Undergoes change of state
Sentience (and/or perception)	Incremental theme
Causing an event or change of state in another participant	Causally affected by another participant
Movement (relative to the position of another participant)	Stationary (relative to the position of another participant)
(Exists independently of the event named by the verb)	(Does not exist independently of the event, or not at all)

**Table 1.1:** Proto-role properties of Dowty (1991, p. 572).

- (38) a. The extremist assassinated the president.  
b. John built a house.

Dowty's focus lies on formulating subject and object selection principles referring to the Proto-Agent and Proto-Patient roles. His argument selection principles are listed in (39) (see Dowty, 1991, p. 576). The relevant principles for two-place subject/direct object verbs are (39a) and (39b): (39a) regulates subject/object selection via a simple counting algorithm, which links the argument bearing most Proto-Agent properties as subject and the one with most Proto-Patient properties as object. This principle suffices to take care of argument linking for semantically clearly asymmetrical predicates, such as in (38) and to explain why the arguments of such verbs are cross-linguistically realised in syntactically transitive constructions. One important characteristic of Dowty's proposal is that subject/object selection is implemented as an inherently *relational* process, since the Proto-Agent and Proto-Patient roles are assigned by *comparing* an argument's number of Proto-Agent and Proto-Patient properties to those of all other arguments of a verb. The absolute number of Proto-Properties a verb entails for a given argument plays no special role in the Proto-role theory, since role identification is a relative process.

- (39) a. ARGUMENT SELECTION PRINCIPLE: In predicates with grammatical subject and object, the argument for which the predicate entails the greatest number of Proto-Agent properties will be lexicalized as the subject of the predicate; the argument having the greatest number of Proto-Patient entailments will be lexicalized as the direct object.
- b. COROLLARY 1: If two arguments of a relation have (approximately) equal numbers of entailed Proto-Agent and Proto-Patient properties, then either or both may be lexicalized as the subject (and similarly for objects).
- c. COROLLARY 2: With a three-place predicate, the nonsubject argument having the greater number of entailed Proto-Patient properties will be lexicalized as the direct object and the nonsubject argument having fewer entailed Proto-Patient

properties will be lexicalized as an oblique or prepositional object (and if two nonsubject arguments have approximately equal numbers of entailed P-Patient properties, either or both may be lexicalized as direct object).

- d. NONDISCRETENESS: Proto-roles, obviously, do not classify arguments exhaustively (some arguments have neither role) or uniquely (some arguments may share the same role) or discretely (some arguments could qualify partially but equally for both proto-roles).

Crucially, Dowty's principles can also account for more variable linking patterns within and across languages with dyadic verbs whose arguments are semantically less clearly distinguished. One well known case discussed by Dowty are what has become known as 'psych' verbs (e.g. Belletti & Rizzi, 1988; Grimshaw, 1990), which fall into two different classes represented by *fear/like* and *frighten/please*, respectively. While the following examples from Levin and Rappaport Hovav (2005, p. 14) are near-paraphrases, the arguments are linked inversely between the (a) and (b) examples: thus, each sentence contains what is usually classified as an 'Experiencer' argument, but it is realised as subject in (40a) and (41a) and as object in (40b) and (41b).

- |      |   |                            |
|------|---|----------------------------|
| (40) | a. My children fear thunderstorms.          | <i>experiencer-subject</i> |
|      | b. Thunderstorms frighten my children.      | <i>experiencer-object</i>  |
| (41) | a. I like this solution to the problem.     | <i>experiencer-subject</i> |
|      | b. This solution to the problem pleases me. | <i>experiencer-object</i>  |

Dowty (1991) points out that one conspicuous characteristic of such psychological predicates is that the Experiencer argument as well as the non-Experiencer argument (often labelled as 'Stimulus') have one Proto-Agent property: the Experiencer bears the entailment of sentience and the stimulus that of causation. On a closer look, however, the psych verbs illustrated by *frighten/please*, however, differ from the *fear/like* psych verbs in further entailing a change of state for the Experiencer object argument; Dowty argues that it is this additional Proto-Patient property which accounts for the fact that *fear/like* type verbs only allow for a stative, non-causative interpretation, while the verbs belonging to the *frighten/please* class are generally ambiguous between such a stative reading and a non-stative, change-of-state interpretation, a difference illustrated by the examples in (42a) and (42b), respectively (from Levin & Rappaport Hovav, 2005, p. 56).

- |      |                                  |                                     |
|------|----------------------------------|-------------------------------------|
| (42) | a. Ghosts frighten me.           | <i>stative</i>                      |
|      | b. The loud noise frightened me. | <i>non-stative, change-of-state</i> |

Thus, the two co-arguments of *fear/like* verbs are characterised by a tie in Proto-properties – both carry one Proto-Agent entailment making the Experiencer and Stimulus argument qualify equally for linking as subject (or object). The Experiencer of *frighten/please* verbs, on the other hand, qualifies as Proto-Patient, since its additional change-of-state entailment present

in non-stative readings creates an asymmetry between the two co-arguments. This asymmetry in Proto-properties can explain why *frighten/please* verbs cross-linguistically behave like prototypical causative verbs such as *break* or *kill* in terms of their argument realisation (the Experiencers are uniformly realised as objects), while *fear/like* verbs exhibit a lot of variability in their linking patterns within a language and across languages (see Croft, 1993).

In a similar manner, Dowty shows how his approach can explain differences between *break*- and *hit*-type verbs (such as that illustrated by example (7) above), how it handles linking with (near) symmetric predicates (cf. examples (26) and (27), for example) as well as alternative argument realisations such as the locative alternation. Further, the second corollary to Dowty's *Argument Selection Principle* (39c) regulates the linking of arguments of three-place verbs using the same counting principle as for two-place verbs without further stipulation. Another advantage inherent to the prototype proposal pointed out by Dowty (1991, p. 578) is the fact that the role rankings which have to be explicitly stipulated with traditional thematic hierarchies emerge naturally with his Proto-roles; as long as a given argument has more Proto-Agent entailments than its co-argument, it will qualify as Proto-Agent, outranking the other argument for subject linking, explaining the priority of Instrument-like arguments over Patient/Theme arguments for subject realisation, for example.

Dowty's approach can thus account for a number of phenomena related to core aspects of argument realisation with only a few intuitive principles and two Proto-roles only. Crucially, his theory gains a lot of flexibility by explicitly treating thematic roles as 'fuzzy' concepts with unclear boundaries, with a given argument of a verb being located on an Agent/Patient *continuum* rather than being discretely classified (39d), allowing the prototype approach to deal with a number of challenging cases in core argument realisation.

While the Proto-role theory as formulated by Dowty has met with success, its limitations have been pointed out by a number of authors (for discussion see Levin & Rappaport Hovav, 2005, section 3.1.1.2). Davis and Koenig (2000) note that Dowty's proposal presupposes that the transitivity of relevant verbs is already known, since the subject/object selection rules assign arguments to one or the other – but as pointed out by Levin and Rappaport Hovav, this assumption is non-trivial. Dowty himself raises the possibility that some of his Proto-properties might be more relevant for argument realisation than others and that his approach may have to be complemented by a weighting of relevant entailments; thus, causes always outrank other arguments for subject selection (cf. Davis & Koenig, 2000; Levin & Rappaport Hovav, 2005) and affectedness plays a prominent role in determining the availability of transitive case frames (Tsunoda, 1985). Other points of criticism raised include that the paired nature of the Proto-Agent and Proto-Patient entailments (e.g. 'causing an event or change of state in another participant' stands in a causal relation to 'undergoes a change of state') stands as an accident in Dowty's approach and that some of his proposed Proto-properties, such as affectedness, can be further simplified and broken down into constituent properties (Grimm, 2011, p. 520). Finally, Levin and Rappaport Hovav (2005, p. 59) point out the somewhat limited scope of Dowty's formulation of the Proto-role theory, since it does not address argument linking phenomena found in morphologically richer languages than English and lacks embedding within a broader theory of grammar.

#### 1.4.2.2 Extensions and adaptations of Dowty's Proto-role approach

A number of researchers have adopted and further developed Dowty's prototype-based linking approach (e.g. Ackerman & Moore, 2001; Blume, 1998, 2000; Davis & Koenig, 2000; Grimm, 2011; Primus, 1999). With the aim to broaden the scope of Dowty's theory, Ackerman and Moore (2001) introduce a new and aspectually based Proto-Patient entailment they term 'bounding entity', for example. Primus (1999) also enlarges Dowty's inventory, and proposes an additional 'Recipient' Proto-role to deal with argument realisation patterns involving dative arguments.

Blume (1998, 2000) makes use of a modified inventory of Proto-properties and combines Proto-role-based linking principles with a subeventual analysis of verbal semantics to address argument linking with two-place verbs which show NOM/DAT rather than the canonical, syntactically transitive NOM/ACC linking in several languages, including German (see the examples in (43) with the verbs *helfen/danken* (*help/thank*)). In her model, Proto-Patient properties are always dependent on Proto-Agent entailments and she derives a semantic transitivity scale based on the assignment of Proto-properties within subevents. Those two-place verbs which license non-canonical NOM/DAT linking are those which exhibit a relatively low degree of transitivity, which Blume takes to be the case when the two arguments are semantically close to each other, i.e. when they do not show an Agent/Patient asymmetry (see section 1.5.2.1 below).

- (43) a. Der        Junge        half        dem        Mann.  
         the.NOM boy.NOM helped the.DAT man.DAT  
         'The boy helped the man.'
- b. Der        Junge        dankte        dem        Mann.  
         the.NOM boy.NOM thanked the.DAT man.DAT  
         'The boy thanked the man.'

Grimm (2011) presents a framework which aims to provide a semantic basis for argument realisation via case marking. He defines a set of Agentivity features, which correlate with Dowty's Proto-Agent entailments; Proto-Patients are taken to be defined via the absence of the very same properties. Table 1.2 lists the Agentive features in the left column and the right columns shows that prototypical Patients are characterised by the *lack* of exactly the same set of entailments. Note that – since all properties are positively defined from the perspective of prototypical Agents – affectedness is here rephrased as 'persistence', which is not taken to be a primitive, but broken down into four separate properties listed in the last four rows of table 1.2 (these will be discussed in section 1.5.1 below).

Grimm (2011) arranges these properties into a hierarchically organised lattice, which represents a structured semantic space in which prototypical Agents are located at the 'high' end of this 'Agentivity lattice' and prototypical Patients at the opposing, hierarchically 'low' end. The embedding of Agentivity features into such a lattice allows to characterise the relationships between different participant types in a principled and intuitive way: the higher/lower



Agentive	Non-Agentive ('Patient')
+ volition	∅ volition
+ sentience	∅ sentience
+ instigation	∅ instigation
+ motion	∅ motion
+ existential persistence (beginning)	∅ existential persistence (beginning)
+ existential persistence (end)	∅ existential persistence (end)
+ qualitative persistence (beginning)	∅ qualitative persistence (beginning)
+ qualitative persistence (end)	∅ qualitative persistence (end)

**Table 1.2:** Grimm's (2011, p. 520) privative Agentivity features defining prototypical Agents – prototypical Patients are defined via the absence of Agentive features.

one moves within the lattice, the higher/lower the degree of Agentivity of the respective argument and the higher/lower its semantic prominence. In addition, topological notions often referred to in defining a semantically based notion of the degree of transitivity such as arguments 'being close semantically' receive a natural definition in terms of the distance of two co-arguments of a dyadic verb within the Agentivity lattice; with verbs denoting canonical transitive events (such as *break/kill*), the two arguments correspond to prototypical Agents and Patients and will be situated at opposing ends of the lattice, while they will be located closer together for verbs which deviate in some way from maximal semantic transitivity (e.g. *hit/fear/see/help*).

Case assignment to arguments is regulated by the assumption that cases 'inhabit' connected regions of the semantic Agentivity lattice (subject to cross-linguistic variation); in nominative/accusative languages, for example, the upper (Proto-Agent) region of the lattice hosts nominative case and the lower (Proto-Patient) one accusative case, whereas the middle regions coincide with oblique cases such as dative. Grimm argues that his model can thus explain linking patterns of three-place NOM/DAT/ACC verbs as well as the non-canonical uses of dative with two-place NOM/DAT verbs investigated by Blume (1998, 2000).

## 1.5 Focussing on affectedness

In this section, I will take a closer look at affectedness as one grammatically relevant aspect of verbal meaning implied in argument linking processes. The following section introduces affectedness as a graded rather than all-or-none phenomenon and section 1.5.2 discusses its role in two correlating linguistic phenomena, transitivity and lexical aspect.

### 1.5.1 Affectedness as a graded phenomenon influencing semantic argument prominence

Affectedness is usually conceived of as the attribute of having undergone some kind of change (e.g. change of state or location) and Beavers (2011, p. 339) provides a list of physical properties which have been taken to represent different types of affectedness for an entity *x*:

- (44) a. *x* changes in some observable property. (*clean/paint/delouse/fix/break x*)  
b. *x* transforms into something else. (*turn/carve/change/transform x into y*)  
c. *x* moves to and stays at some location. (*move/push/angle/roll x into y*)  
d. *x* is physically impinged. (*hit/kick/punch/rub/slap/wipe/scrub/sweep x*)  
e. *x* goes out of existence. (*delete/eat/consume/reduce/devour x*)  
f. *x* comes into existence. (*build/design/construct/create/fashion x*)

The use of primitive semantic predicates such as *BECOME*, which serve to encode such changes in some approaches (see example (35), for example), may suggest that affectedness involves only a binary distinction, i.e. change or no change. However, it is often acknowledged that affectedness is actually a graded phenomenon, as illustrated by the sentences in (45), which are based upon Beavers (2011, p. 336) and Beavers (2013, p. 688) and in which the object *the apple* is understood to be increasingly less affected (my emphasis)<sup>5</sup>:

- (45) a. John devoured the apple. (*Apple is completely gone*)  
b. John cut the apple. (*Apple cut, not necessarily to a particular degree*)  
c. John hit the apple. (*Apple impinged, not necessarily affected*)  
d. John touched/saw the apple. (*Apple not necessarily impinged*)

As noted by Beavers (2011), degrees of affectedness have played a role in the explanation of a number of linguistic phenomena (such as transitivity – see 1.5.2.1 below), but at the same time affectedness as a graded concept has often escaped an exact definition. While higher affectedness contributes to increased transitivity of clauses in the study of factors defining transitivity of Hopper and Thompson (1980), in that “[t]he degree to which an action is transferred to a Patient is a function of how completely that Patient is AFFECTED” (pp. 252–253, their emphasis), for example, Beavers (2011, p. 336) points out that they actually fail to define this degree of transfer. Næss (2004, p. 1202) remarks that “affectedness can take many different forms and is not easily quantified” and proposes that an event participant is more affected if more of the respective entity is affected (e.g. the whole object is broken, rather than just a part of it) and if the respective effects are perceived as more

<sup>5</sup>Note that Beavers (2013) provides *see* as an example verb for the lowest affectedness level, while Beavers (2011, p. 336) gives *touch* as a representative verb and adds “Apple **manipulated**, not necessarily impinged” as a description (my emphasis). Since I am using both verbs as examples in (45), I am leaving out the property of being manipulated, since it does not fit well with the verb *see*.

*salient* by humans – thus, *breaking a vase* is supposed to be perceived as implying a less ‘significant/dramatic’ kind of affectedness than *killing John*, for example. Beavers (2011, p. 357) points out that the problem with this view of degrees of affectedness is to find a way of formalization if a concept such as ‘salience of effect for humans’ has any actual linguistic consequences.

Another way of approaching degrees of affectedness is illustrated by the work of Grimm (2011), who arranges the prototype features listed in table 1.2 above in a structured lattice, as discussed in the previous section. He splits up ‘persistence’ – which is the Agentive counterpart of affectedness – into ‘existential persistence’ and ‘qualitative persistence’: the former implies that a participant’s “essence remains the same in that it neither comes into or goes out of existence during the course of the event/state”, while the latter means that “none of its qualities undergo any change” and “covers any other change besides an existential one, including changes of quality, quantity or location” (Grimm, 2011, p. 521). Thus, Grimm’s definition of degrees of affectedness is intuitively more concrete than the one of Næss, for example.

One more recent approach to degrees of affectedness I will take a somewhat closer look at is the one provided by Beavers (2010, 2011). Beavers groups verbal predicates into four different classes implying different degrees of affectedness for their direct objects. The four groups of verbs are illustrated by the examples in (45) above: a verb like *devour* (45a) entails that the apple is *completely* gone (i.e. affected), while *cut* in (45b) means that it is cut (i.e. affected), but does not imply that it is cut to a *specific* degree; with *hit* (45c), on the other hand, the apple is impinged, but not necessarily affected in any way and a verb like *see* or *touch* (45d) does not even imply impingement (let alone affectedness) of the apple.

Crucially, these four groups of verbs can be distinguished by specific linguistic tests (see Beavers, 2011, section 2); the verbs in (45a)–(45c) all have objects which qualify as ‘force recipients’, i.e. they might change as the result of bearing the brunt of the force of the action denoted by the predicate (see Rappaport Hovav & Levin, 2001). The relevant test identifying such force recipient objects is *What happened to X is Y* (see also Cruse, 1973, for example). Thus, verbs taking force recipient objects (46a) pass the test, while (46b) is odd.

- (46) a. What happened to the apple is that John devoured/cut/hit it.
- b. # What happened to the apple is that John saw it.

The subset of force recipients which actually entail some change of state can be determined using the *but nothing is different about X* entailment test which denies any change in the participant linked as the object: while verbs such as *devour/cut* are grouped together by this test as in (47a), *hit* and *see* in (47b) fail the test, indicating that these verbs do not *necessarily* entail any change.

- (47) a. # John devoured/cut the apple, but nothing is different about it.
- b. John hit/saw the apple, but nothing is different about it.

Diagnostics	<i>devour x</i>	<i>cut x</i>	<i>hit x</i>	<i>see x</i>
Telic	✓	✗	✗	✗
Entails change	✓	✓	✗	✗
<i>What happened to X is Y</i>	✓	✓	✓	✗

**Table 1.3:** Verbal predicates grouped into four different affectedness classes via a subset relation (see Beavers, 2011) ; the table is based upon the one presented by Beavers (2013, p. 689).

And finally, only a subset of these obligatorily change-entailing predicates are telic, i.e. they imply a natural and inherent endpoint for the event denoted (see Dowty, 1979; Vendler, 1957, for example; also see the discussion of the relation between affectedness and lexical aspect in section 1.5.2.2 below). One standard test for telicity is modification by time frame adverbials such as *in X minutes/days/hours* and time span adverbials such as *for X minutes/days/hours*; while acceptability may depend upon a number of factors, on the most salient, intended readings the former are best compatible with telic verbs, while the latter are most natural with atelic verbs. Using this diagnostic, only the verbs belonging to the *devour* class qualify as telic (48a), while all others are atelic (48b).

- (48) a. John devoured the apple in/?for five minutes.  
b. John cut/hit/saw the apple for/#in five minutes.

While Beavers (2011) suggests additional tests, these suffice to show that the predicate types listed in (45) can be grouped into a subset relation: as table 1.3 illustrates, verbs passing  $n$  of the above tests are a subset of those satisfying  $n-1$  tests.

Beavers (2011) builds upon these findings to construct a principled model of affectedness which is based upon a scalar approach to change of state as worked out by Hay, Kennedy, and Levin (1999) and Kennedy and Levin (2008); within this model, a predicate of change  $\varphi$  is analysed as a relation between three entities, an event  $e$ , a theme  $x$  that changes in the course of  $e$  and a scale  $s$  defining the change which  $x$  undergoes in  $e$ . Within this context, a scale  $s$  represents an ordered set of degrees indicating the extent of having a specific property and a change of state is modelled as theme  $x$  transitioning from one degree of possessing a given property to another degree. One advantage of such a scalar approach to degrees of affectedness is that the scale can be separated from the type of change denoted by a verb, which allows to address any kind of change in the same way while abstracting away from real-world properties of the change denoted by the predicate: thus, “[d]ifferent real world changes are distinguished by the *type* of the scale (e.g. scales of cleanliness, color, dirtiness, volume, position, etc.), but can be compared to one another in terms of the *structure* of the scale, regardless of type” (Beavers, 2011, pp. 337–338; emphasis in the original). As a consequence, all different types of change – such as creation/consumption, change of state or change of location – can be modelled uniformly and formally treated in the same way.

The upshot of Beavers' scalar model of affectedness is that degrees of affectedness can now be defined in terms of how *specific* a given verb is about the progress of the theme  $x$  on the scale  $s$ , with each of the predicate classes illustrated in example (45) and grouped by means of a subset relation (see table 1.3) being less specific about the final state of theme  $x$  on  $s$  (see Beavers, 2011, section 5): predicates at the highest affectedness level (e.g. *devour* – see (45a)) entail a *quantised* change, i.e. the theme  $x$  transitions to a *specific* state  $g_\phi$  on the respective scale, and they pass all of the three tests discussed above. Similarly, verbs of the *cut* class (45b) entail a change for their theme  $x$ , which differs from a quantised change, however, in that  $x$  only undergoes a *non-quantised* change, i.e. *some* (non-specific) result state  $g$  different from its initial state holds for theme  $x$ . Since their direct objects are force recipients and actually undergo some change, these predicates pass the two respective tests, but fail telicity tests since the events denoted lack inherent temporal endpoints. The direct objects of impact/surface contact predicates such as *hit* (45c) qualify as pure force recipients and only pass the *What happened to X is Y* test; since the respective participants *may* change, but do not necessarily undergo change, Beavers refers to the verbs of this affectedness level as *potential change* predicates and models such a potential for change by assuming that a latent scale argument exists for these verbs without any transition of  $x$  along the scale being entailed. Finally, the objects of verbs of the *see*-type (45d) at the lowest level of affectedness fail all the tests – as non-force recipients, they are assumed to be associated with an event without a scale argument. Beavers (2011, p. 358) provides the following formal definitions of these four degrees of affectedness for a theme  $x$  associated with a given verb  $\phi$ :

- (49) a.  $x$  undergoes a **quantized** change iff  $\phi \rightarrow \exists e \exists s [\text{result}'(x, s, g_\phi, e)]$   
 (e.g. accomplishments/achievements: *break, shatter, destroy, devour x*)
- b.  $x$  undergoes a **non-quantized** change iff  $\phi \rightarrow \exists e \exists s \exists g [\text{result}'(x, s, g, e)]$   
 (e.g. degree achievements: *widen, cool, lengthen, cut, slice x*)
- c.  $x$  has **potential** for change iff  $\phi \rightarrow \exists e \exists s \exists \theta [\theta(x, s, e)]$   
 (e.g. surface contact/impact: *wipe, scrub, rub, punch, hit, kick, slap x*)
- d.  $x$  is **unspecified** for change iff  $\phi \rightarrow \exists e \exists \theta' [\theta'(x, e)]$   
 (e.g. other activities/states: *see, laugh at, smell, follow, ponder, ogle x*)

As noted above, different types of change can be modelled in parallel by using different types of scales; for motion predicates entailing a change of location for the theme  $x$ , the scale  $s$  represents a *location* scale (i.e. a path) as in (50a), for change of state verbs it is a property scale (50b) and *creation/consumption* predicates are associated with an extent scale as in (50c). Note that all decompositions in (50) (which corresponds to example (49) of Beavers, 2011, p. 352) are formally parallel, with the first conjunct representing the respective process and the second one the entailed result state of  $x$  – thus, differences in real world concepts associated with different types of change (i.e. differences in semantic content) are set aside and commonalities in semantic structure are captured.

- (50) a. John walked to the cafe. (scale  $s$  of *position* of John)  
 $\exists e \exists s [\text{walk}'(\text{john}, s, e) \wedge \text{result}'(\text{john}, s, \text{cafe}, e)]$   
 b. John wiped the table clean. (scale  $s$  of *cleanliness* of the table)  
 $\exists e \exists s [\text{wipe}'(\text{john}, s, \text{table}, e) \wedge \text{result}'(\text{table}, s, \text{clean}, e)]$   
 c. John ate the apple. (scale  $s$  of *volume/existence* of the apple)  
 $\exists e \exists s [\text{eat}'(\text{john}, s, \text{apple}, e) \wedge \text{result}'(\text{apple}, s, 0, e)]$

The four degrees of affectedness make up the implicational *Affectedness Hierarchy* in (51) (Beavers, 2011, p. 359); being implicational, if  $x$  is associated with a given degree  $n$  of affectedness on the hierarchy, it also bears all degrees lower than (i.e. to the right of)  $n$ . Conceptually, Beavers thus recasts degrees of affectedness in terms of the strength of truth conditions about the effect a predicate entails for its participant realised as the direct object, here encoded as the specificity of the transition of a theme  $x$  on a scale  $s$  and tested via the diagnostics introduced above. A higher degree of affectedness  $n+1$  corresponds to a monotonic strengthening of the truth conditions associated with degree  $n$ .

- (51) For all  $x, \phi, e$ , *quantised*  $\rightarrow$  *non-quantised*  $\rightarrow$  *potential*  $\rightarrow$  *unspecified*

Beavers (2010) builds upon this scalar model of affectedness and explores its implications for the semantic basis of argument realisation. He argues that approaches to argument realisation which solely rely upon predicate/event structure decompositions (see section 1.4.1) cannot naturally account for some phenomena and require recourse to ad-hoc decompositions. To illustrate this point, Beavers focusses on direct/oblique alternations in which one argument (or more) vary between being encoded as a direct or as an oblique object. A case in point is the *conative alternation* exemplified in (52)–(54) (all examples from Beavers, 2010, p. 830), in which the object is either linked as direct object or as a PP headed by *at*. As Beavers (2010, p. 830) notes, the semantic difference between the direct object variants (a) and their oblique counterparts (b) is not generally a matter of presence or absence of change of state, but the induced contrast depends upon the respective verb used: with a verb entailing quantised change for its direct object, such as *eat/drink* in (52a), the oblique conative variant (52b) also entails change, but of the non-quantised type. For non-quantised change predicates such as *cut/slice* (53a), the conative encoding in (53b) entails a weaker degree of affectedness, i.e. potential for change; when the direct object realisation entails potential for change (e.g. for *hit/smack* as in (54a)), the conative variant (54b) is unspecified for change.

- (52) *Affectedness demotion: from quantised to non-quantised*  
 a. Marie ate her cake/drank her wine. (all of cake/wine consumed)  
 b. Marie ate at her cake/drank at her wine. (at least some cake/wine consumed)
- (53) *Affectedness demotion: from non-quantised to potential change*  
 a. Marie cut/sliced/scratched/slashed the rope. (the rope is cut)

- b. Marie cut/sliced/scratched/slashed at the rope. (*the rope may or may not be cut*)
- (54) *Affectedness demotion: from potential to unspecified change*
- a. Marie hit/slapped/smacked/whacked Defarge. (*Defarge hit, not necessarily affected*)
  - b. Marie hit/slapped/smacked/whacked at Defarge. (*Defarge not necessarily even hit*)

Thus, in Beavers's approach to affectedness, the conative alternation can be analysed as a systematic demotion of the degree of affectedness entailed for the theme *x*, shifting the type of change involved given the direct object variant of a specific verb down by one along the Affectedness Hierarchy in (51). Beavers (2010) argues that these systematic patterns can be explained in a straightforward way if degrees of affectedness are interpreted as semantic prominence; on this view, the semantic prominence of an argument depends upon the truth conditions associated with it and stronger truth conditions correspond to higher semantic prominence. Since arguments located high on the Affectedness Hierarchy have stronger truth conditions than those of lower levels, arguments which are more specific about the change undergone are semantically more prominent than those which are less specific. Beavers (2010, p. 831) formulates the *Morphosyntactic Alignment Principle* (MAP) in (55)<sup>6</sup>, which allows to treat direct/oblique alternations simply as a case of prominence preservation (see section 1.3.2), given independently motivated morphosyntactic prominence hierarchies such as in (32) and (33) (repeated below). According to this principle, stronger truth conditions with respect to affectedness entail a more prominent morphosyntactic encoding for the respective argument.

- (55) Morphosyntactic Alignment Principle (MAP) (version 1): In direct/oblique alternations the direct realization of an alternating participant has as strong or monotonically stronger truth conditions associated with it than its corresponding oblique realization.

(32) subject > object > indirect object > oblique

(33) nominative > accusative > dative > oblique cases

Note that the MAP regulates realisation of the *same argument across clauses*, rather than the linking of *co-arguments within one and the same clause* (see also Ackerman & Moore, 2001, who propose a similar 'Paradigmatic Argument Selection Principle' for such cases) as, for example, required for subject/object selection. However, Beavers (2010, pp. 846–847) posits two additional principles which handle clause-internal linking of co-arguments and are variants of Dowty's (1991) Argument Selection Principle. Since Beavers formulates his argument realisation principles within Dowty's entailment-based theory of lexical meaning and

<sup>6</sup>The version of the MAP given here is a preliminary one, but suffices for the present purposes – the final version of the MAP is provided on page 848 of Beavers (2010).

Dowty's entailments	Beaver's entailments
a. incremental theme	totally traversed
b. holistic theme	quantised change
c. undergoes change of state	non-quantised change potential change
d. stationary (relative to movement of another participant)	potential change?
e. causally affected by another participant	(non-)quantised change?
(f. does not exist independently of the event, or not at all)	((non-)quantised change)

**Table 1.4:** Possible correspondences between the Proto-Patient entailments originally proposed by Dowty (1991) and those of Beavers (2010).

argument linking, the different degrees of affectedness can be interpreted as analogues to Dowty's Proto-Patient entailments. Table 1.4 (from Beavers, 2010, p. 853) lists Dowty's Proto-Patient properties together with the corresponding degrees of affectedness proposed by Beavers. Note that 'totally traversed' is an additional entailment belonging to a 'Traversal Hierarchy' Beavers posits for motion predicates with path scales, which works parallel to the Affectedness Hierarchy and "although coverage of the path is not itself a type of affectedness, it is the flipside of the degree of affectedness of the theme, which is a relation between the theme and the scale" (Beavers, 2010, p. 853).

Beavers (2010, p. 852) acknowledges that affectedness may not be the only factor in determining objecthood, which is to be expected given that Proto-roles are supposed to be made up of semantically heterogeneous sets of entailments. And while his approach to argument linking based upon strength of truth conditions represents an attempt at such an alternative view of the syntax/semantics interaction implied in argument linking, it provides an interesting link to Dowty's prototype theory of argument realisation and can parsimoniously account for a number of different patterns, such as argument alternations of the same argument across clauses (such as the conative and locative alternation – for the latter, see for example Beavers, 2010, pp. 839–841) and linking of co-arguments within a clause.

Beavers also discusses the relation between his proposal and event structure approaches based on predicate decompositions and notes that the two are actually complementary: since causers outrank all other arguments for subject linking and temporally and causally precede all others (see Croft, 1998, among others), "decompositions of at least the form [ $\varphi$  CAUSE  $\psi$ ] may well be motivated. [...] Thus we can say that the MAP fills in a gap in explanation where subevent structure fails, namely what happens below CAUSE in [ $\varphi$  CAUSE  $\psi$ ]" Beavers (2010, pp. 857–858).



## 1.5.2 Correlating phenomena

### 1.5.2.1 Transitivity

Affectedness has been known to be correlated with the phenomenon of transitivity (e.g. Beavers, 2010, 2011; Hopper & Thompson, 1980; Levin, 1999; Malchukov, 2005; Tsunoda, 1981, 1985). While transitivity is most commonly conceived of in syntactic terms, there are actually a number of (partially) separable notions of transitivity, which are reviewed by Blume (2000, section 6.2); on the syntactic view, a verb is traditionally defined as transitive if it subcategorises for two (or more) arguments which are morphosyntactically realised in an unmarked way. In case marking languages of the nominative/accusative type, such as German, dyadic verbs with nominative subjects and accusative objects (and in ergative languages absolutive/ergative verbs) thus qualify as syntactically transitive. In German, an additional diagnostic for syntactic transitivity is the ability to form a verbal passive with the auxiliary *werden*, linking the accusative object as the nominative subject.

According to these more traditional criteria, a verb like *waschen/wash* (56a) counts as syntactically transitive; while a verb like *enthalten/contain* (56b) meets the first criterion (nominative subject/accusative object), building a *werden*-passive is not possible and can thus be seen as 'pseudo-transitive'. Further, verbs such as *helfen/help* and *gefallen/please* in (56c) and (56d) would count as intransitive due to their nominative/dative linking pattern and verbs such as *schlafen/sleep* because they are monadic (examples are from Blume, 2000, p. 177).

(56) a. *Syntactically transitive:*

Anna        wäscht den        Hund.  
Anna.NOM washes the.ACC dog.ACC  
'Anna washes the dog.'

b. *Syntactically pseudo-transitive:*

Die        Kiste        enthält Bücher.  
the.NOM box.Nom contains books.ACC  
'The box contains books.'

*Syntactically intransitive:*

- c. Albert        hilft dem        Kunden.  
Albert.NOM helps the.DAT customer.DAT  
'Albert helps the customer.'
- d. Der        Aufsatz        gefällt dem        Lehrer.  
the.NOM essay.Nom pleases the.DAT teacher.DAT  
'The essay pleases the teacher.'
- e. Sara        schläft.  
Sara.NOM sleeps  
'Sara is sleeping.'

	High	Low
a. Participants	2 or more participants, A and O	1 participant
b. Kinesis	action	non-action
c. Aspect	telic	atelic
d. Punctuality	punctual	non-punctual
e. Volitionality	volitional	non-volitional
f. Affirmation	affirmative	negative
g. Mode	realis	irrealis
h. Agency	A high in potency	A low in potency
i. Affectedness of O	O totally affected	O not affected
j. Individuation of O	O highly individuated	O non-individuated

**Table 1.5:** Transitivity components of Hopper and Thompson (1980, p. 252) – ‘A’ stands for Agent and ‘O’ for Object participants in a clause with two realised arguments.

On a complementary view, transitivity can be defined via semantic rather than syntactic properties of verbs; one of the most widely known studies in this spirit is that of Hopper and Thompson (1980), who view transitivity as a multi-factorial phenomenon composed of a number of semantic determinants, which are listed in table 1.5. In their view, transitivity is a gradable property which is associated with *whole clauses*, rather than lexically determined aspects of verb meaning only. Thus, as pointed out by Blume (2000, section 6.2.2), the majority of Hopper and Thompson’s features are context dependent, i.e. not related to inherently lexical semantic aspects of verbs: among the context dependent aspects are properties such as ‘affirmation’, ‘mode’ and ‘individuation of O’, but only ‘action’, ‘volition’ and ‘high in potency’ can be identified as lexically determined Proto-Agent (belonging to ‘kinesis’, ‘volitionality’ and ‘agency’, respectively) and ‘totally affected’ as verb-related Proto-Patient properties (Blume, 2000, p. 183–185).

As outlined briefly in section 1.4.2.2, Blume (2000) combines a modified version of Dowty’s prototype-based theory of argument realisation with a subeventual analysis of verbal semantics to extend the scope of the prototype approach to linking and to account for marked linking patterns such as NOM/DAT verbs in German, which are exemplified in (43). Blume separates syntactic from semantic transitivity and takes both to be graded concepts. As pointed out above, in her view – and contra Hopper and Thompson – semantic transitivity depends solely upon lexically determined properties and is defined by the ‘semantic distance’ between co-arguments of a verb: it is maximal with core examples of transitive verbs such as Agent/Patient taking verbs (e.g. *kill/break*) and reduced in cases with a distribution of Proto-roles which does not result in a clear semantic asymmetry between co-arguments (see Blume, 2000, section 6.2.2).

Blume notes that although the degree of syntactic and semantic transitivity often correlate, they are independent to a certain degree. Thus, a high degree of syntactic transitivity does not necessarily result in an equally high degree of semantic transitivity, as illustrated by the examples in (57) in which the arguments are all linked with the unmarked NOM/ACC pattern, but none of which exhibit maximal semantic transitivity. On the other hand, “a semantically strongly transitive verb is always syntactically strongly transitive – or put differently: a verb whose syntactic transitivity is restricted is never semantically strongly transitive” (Blume, 2000, p. 197, my translation of the German original).

- (57) a. Luk           bewundert seine    Schwester.  
           Luk.NOM admires    his.ACC sister.ACC  
           ‘Luk admires his sister.’
- b. Sie           sah ihn.  
           she.NOM saw him.ACC  
           ‘She saw him.’
- c. Die           Familie       unterstützt den       Sohn.  
           the.NOM family.Nom supports   the.ACC son.ACC  
           ‘The family supports the son.’

Remember that Grimm (2011, discussed in 1.4.2.2) similarly defines the degree of a verb’s semantic transitivity as semantic proximity between co-arguments, but takes the semantic distance to be a natural by-product of the position of a verb’s arguments within his Agentivity lattice. What Grimm’s and Blume’s approaches have in common is the ‘dipolar’ basis of their idea of semantic transitivity, in the sense that the degree of prototypicality of *all* arguments is relevant for assessing the semantic distance between a verb’s arguments – and thus also its degree of semantic transitivity (see also Malchukov, 2005; Testelec, 1998, for related proposals).

Other researchers have concentrated on semantic properties of the object alone for assessing a verb’s degree of affectedness. In his discussion of Hopper and Thompson’s multifactorial approach to transitivity, Tsunoda (1985, p. 386) points out that some of the proposed parameters co-vary strongly, but others don’t: while volitionality and agency tend to co-vary, the relation is much weaker in other combinations and does not exist between volitionality/agency on the one hand and affectedness on the other hand. Tsunoda (p. 386) argues that “[t]he ten parameters are not equally relevant to a given morphosyntactic property of transitivity. Thus, in manifesting a transitive case frame, (I) Affectedness is crucial, but (E) Volitionality and (H) Agency appear to be irrelevant”. He concludes that affectedness needs to be further refined to better classify a verb with respect to its degree of affectedness and presents his version of an affectedness scale.

This focus on properties of the object is shared by Levin (1999), differentiating between ‘core transitive verbs’ (CTVs) and ‘non-core transitive verbs’ (NCTVs). She takes the CTVs

to be a subset of the syntactically transitive verbs of a language which conform to the prototypical semantic situation where an Agent acts on and affects a Patient. Formally, in her decompositional event structure framework (see section 1.4.1 above) these correspond to causative event templates, i.e. the template associated with accomplishments given in example (35d) and repeated below. The crucial difference between such a causative event structure template and others lies in the complexity: causative accomplishments (35d) have a complex structural representation in the sense that they consist of two subevents with one participant/argument each, while the simpler templates in (35a)–(35c) are made up of a single subevent with one associated argument only.

- |      |    |  |                       |
|------|----|--|-----------------------|
| (35) | a. | $[x \text{ ACT}_{\langle \text{MANNER} \rangle}]$  | <i>activity</i>       |
|      | b. | $[x \langle \text{STATE} \rangle]$   | <i>state</i>          |
|      | c. | $[\text{BECOME } [x \langle \text{STATE} \rangle]]$  | <i>achievement</i>    |
|      | d. | $[[x \text{ ACT}_{\langle \text{MANNER} \rangle}] \text{ CAUSE } [\text{BECOME } [y \langle \text{STATE} \rangle]]]$ | <i>accomplishment</i> |
|      | e. | $[x \text{ CAUSE } [\text{BECOME } [y \langle \text{STATE} \rangle]]]$   | <i>accomplishment</i> |

According to Levin, this difference between templates in structural complexity leads to distinct licensing mechanisms for the direct objects of CTVs and NCTVs: the causative event structure template features two argument variables, so each of the participants of a CTV can be linked to one of them, rendering them ‘structure arguments’. NCTVs also syntactically realise two arguments, but are taken to be associated with simple event structure which only license one structural argument position for the subject – the object of NCTVs is thus left to be licensed by conceptual properties of the verb’s root/constant and qualifies as ‘pure constant argument’ not rooted in an event structure template (see Grimshaw, 2005b, for a very similar proposal). (58a) provides an example of a one-argument activity verb such as *run* and (58b) illustrates a two-argument activity verb like *sweep*: for verbs like *sweep*, the object is assumed to be linked as a pure constant argument, which is indicated by the underlined *y* participant in the event structure template for *sweep*. Levin argues that one of a number of linguistic manifestations of this difference in linking processes is the higher number of argument expression options for NCTVs, including the optionality of expressing their object, an option which is less readily available for objects of CTVs, as illustrated by (58c) and (58d); since causative event structure templates are associated with two structure participants, both of them have to be realised syntactically, rendering their objects obligatory.

- |      |    |  |   |
|------|----|--|---|
| (58) | a. | John ran.  | <i>One-argument activity verb:</i> $[x \text{ ACT}_{\langle \text{RUN} \rangle}]$                 |
|      | b. | John swept the floor.  | <i>Two-argument activity verb:</i> $[x \text{ ACT}_{\langle \text{SWEEP} \rangle} \underline{y}]$ |
|      | c. | John swept this morning.                                     |   |
|      | d. | John broke the vase this morning./* John broke this morning. |   |

The importance of the degree of affectedness entailed for a syntactically transitive verb in determining its underlying semantic transitivity is also stressed by Beavers (2011, p. 362). He argues that his Affectedness Hierarchy can be interpreted as ordering verbs in the order

of their degree of transitivity, thus providing an independently motivated reinterpretation of transitivity scales formulated in terms of semantic properties of the object (e.g. Malchukov, 2005; Tsunoda, 1985) in terms of strength of truth conditions about the entailed effect on the respective event participant.

### 1.5.2.2 Lexical aspect

Another aspect of verbal semantics which is closely related to affectedness is lexical aspect, which subsumes a number of facets of verbal meaning tied to the inherent temporal properties of verbs, such as whether a verb describes a durative or punctual event (e.g. *run/sleep/paint* vs. *burst/recognise/notice*). The aspectual property most relevant for discussing the relation between affectedness (as implying a change of state) is *telicity*: telic predicates are verbs which involve an endpoint while atelic verbs lack such a terminal point (see Filip, 2011, for a recent overview of facets of lexical aspect). As mentioned in section 1.5.1, telicity of a predicate can be tested for by means of time frame and time span adverbials (*in X minutes/days/hours* and *for X minutes/days/hours*, respectively), with the former being most naturally compatible with telic (see example (48a) repeated below) and the latter with atelic predicates (48b).

- (48) a. John devoured the apple *in/?for* five minutes.  
       b. John cut/hit/saw the apple *for/#in* five minutes.

On some accounts these lexical aspectual properties – and especially telicity – are a central component for argument linking at the syntax/semantics interface. One of the most prominent of these approaches is proposed by Tenny (1992, 1994), who formulates the *Aspectual Interface Hypothesis* stating that “[t]he mapping between thematic structure and syntactic argument structure is governed by aspectual properties” (Tenny, 1992, p. 2). On her view, affectedness can be reduced to the aspectual properties of ‘delimiting’ and ‘measuring out’ an event; while delimiting corresponds to telicity in implying an endpoint in time, ‘measuring out’ refers to “uniform and consistent change, such as change along a scale” (Tenny, 1992, pp. 4–5). Direct objects which undergo a change of state in the event denoted by the respective verb are taken to measure out events: in (59a), the direct objects *an apple/a poem* measure out the respective event through their volume/spatial extent and in (59b), *the shirt/table* do the same via a relevant semantic property, i.e. ‘colourfulness’ and ‘cleanliness’.

- (59) a. eat an apple/translate a poem  
       b. bleach the shirt/clean the table

With such verbs, their direct objects both measure out and delimit the event; with others, such as verbs of motion like *push*, the direct object does not delimit it, as illustrated by the strangeness of a time span adverbial in (60a). Nevertheless, the event can still be delimited by addition of a goal phrase as in (60b) (Tenny, 1992, p. 6). However, the direct object *the cart* is still the argument undergoing the change of location and thus still measures out the event through its location.

- (60) a. push the cart (#in an hour/for an hour)  
 b. push the cart to New York (in an hour/#for an hour)

Tenny takes such patterns to indicate a special relation between direct objecthood and aspectual properties and argues that “[t]he verb’s direct internal argument may be thought of as being converted into a function of time at some level of semantic representation” (Tenny, 1992, p. 4). On this view, the focus in the discussion of affectedness is shifted onto the temporal process underlying a change of state and “[a]n affected argument can be more adequately described in aspectual terms” (Tenny, 1992, p. 8).

Beavers (2011), on the other hand, acknowledges that affectedness and lexical aspect stand in a tight correlation to each other, but argues that the relationship of these two notions is complex and that affectedness cannot be reduced to telicity and measuring out, contra Tenny (1992, 1994). To illustrate his arguments, I will first exemplify a well known fact about telicity; consider the two sentences in (61), which differ only in the referential properties of the direct object – being a mass noun, *wine* in (61a) has a non-quantised reference (i.e. it does not specify a specific quantity of wine consumed), but *a glass of wine* in (61b) refers to a quantised amount of wine. With a quantised object, a verb of consumption such as *drink* yields a telic verbal expression (61b), but with a non-quantised object, it becomes atelic (61a). With other verbs, such as perception verbs like *see*, the referential character of the object does not make a difference: whether quantised (62b) or not (62a), the expression is atelic (both examples are from Krifka, 1992, pp. 30–31).

- (61) a. John drank wine (for an hour/#in an hour).  
 b. John drank a glass of wine (#for an hour/in an hour).  
 (62) a. John saw a zebra (for an hour/#in an hour).  
 b. John saw zebras (for an hour/#in an hour).

Dowty (1991) argues that the direct objects of verbs of consumption/creation (and others) carry the entailment of *Incremental Theme* and follows Krifka (1989, 1992, 1998) in assuming that Incremental Theme objects stand in a homomorphic relation to the event denoted by the verb, i.e. that parts of the entity denoted by an Incremental Theme are mapped onto parts of the event referred to by the respective verb. One crucial consequence of such an approach is that “with certain thematic relations, the reference properties of the syntactic arguments carry over to the reference properties of the complex construction” (Krifka, 1992, p. 38). Informally speaking, with a quantised Incremental Theme its boundedness carries over to the whole verbal expression, imposing an endpoint on it and rendering it telic; a non-quantised (or ‘cumulative’ in Krifka’s terms) Incremental Theme yields an atelic complex verbal predicate.

Beavers (2011) uses these diagnostics and interactions of properties of Incremental Themes with aspectual characteristics of verbal expressions to question the reduction of affectedness to aspectual notions. One argument he brings forward are ‘double telicity effects’, i.e. cases in which the telicity of an expression is not only determined by the referential properties of

the direct object, but jointly by two separate phrases. While a similar point can be made with creation/consumption verbs (see Beavers's example (41) on p. 349), (63) illustrates this effect with *dim*, a verb entailing a change of a property of the object. Here, a definite, specific object like *the lights* AND a specific result phrase like *half dim* together induce telicity (63a), but if either the object appears as a bare plural (63b) or the result state is only vaguely specified (63c), the verbal predicate is atelic.

- (63) a. Bill dimmed the lights half dim in/?for five minutes.  
 b. Bill dimmed lights half dim ??in/for five minutes.  
 c. Bill dimmed the lights dimmer and dimmer ??in/for five minutes.

Beavers (2011, p. 356) thus concludes that these patterns lend support for his model of affectedness, which incorporates a theme as well as a scale which measures the change (see section 1.5.1):

"[T]elicity is not about the entire theme being affected, nor is it about traversal of an entire scale. Telicity is instead a property a predicate has when all of the theme crosses all of the scale, an inherently relational definition. [...] In summary, the scalar analysis I adopt unifies all types of changes under a single analysis and recognizes two entities, a theme and a scale of change, which stand in a mutually constraining relation to each other and the event. This provides a natural way to capture the complex correlation of change and aspect."

On this view, change is a linguistic primitive on its own and not reducible to lexical aspect<sup>7</sup>, a position shared by Rappaport Hovav and Levin (2005), who base their conclusion mainly upon the different argument expression patterns of core examples of change of state verbs (the CTVs of Levin, 1999) and other classes of verbs. Thus, while the object of verbs belonging to other verb classes may be left unspecified, the direct objects of change of state verbs must be expressed syntactically (see also example (58) above) as illustrated by (64a). The more restricted argument realisation options of change of state verbs relative to other verb types is also illustrated by their incompatibility with non-subcategorised NP resultative phrases such as in (65) (all examples are from Rappaport Hovav & Levin, 2005, p. 276).

- (64) a. \* Pat broke/dimmed.  
 (65) a. \* My kids broke me into the poorhouse.  
 b. \* The stagehand dimmed the scene dark.

As Rappaport Hovav and Levin (2005) point out, while change of state verbs such as *break* and *dim* show common argument realisation patterns, these two verbs actually differ in their aspectual properties; while *break* is telic, the degree achievement *dim* is either telic or atelic,

<sup>7</sup>Note that Beavers (2013) actually argues that the standard aspectual verb classes can be naturally derived from his scalar model of change as worked out in Beavers (2010, 2011).

depending on the context. Further, *break* is a punctual predicate, but *dim* is a durative one. These and similar observations lead Rappaport Hovav and Levin to a conclusion in line with Beavers's: the argument realisation options of change of state verbs are governed by a lexicalised, non-aspectual property, which is the property of undergoing a change, in other words their status as affected arguments.

## 1.6 Summary

This chapter has introduced some basic questions and concepts related to the phenomenon of argument linking and has discussed the role of affectedness as a central semantic component of the syntax/semantics interface, at which the mapping from verbal semantics to syntactic structure takes place, in more detail. A review of the traditional approach to argument linking has shown that thematic roles and thematic hierarchies suffer from a number of drawbacks which alternative accounts involving predicate decompositions and Proto-roles (e.g. Dowty, 1991) aim to avoid. A promising perspective on affectedness as a graded phenomenon is provided by Beavers (2010, 2011); importantly, his scalar approach to affectedness does not only offer a solid basis for defining degrees of affectedness, but also opens up direct links to theories of argument linking making use of Proto-roles as well as to work on semantic determinants of transitivity. It further allows to refine the nature of the relation between affectedness and aspectual notions like telicity and suggests that affectedness is in fact a linguistic primitive which is not reducible to lexical aspect. The following chapter builds upon the concepts and approaches discussed above and presents a first experimental investigation of the role of affectedness in the mapping from predicate meaning to syntactic structure.



## Chapter 2

# Probing affectedness: a detour via argument linking in deverbal nominalisations

The current chapter presents a first experimental exploration of possible effects of affectedness with processes of argument linking. Its main aims are to introduce, justify and validate the basic experimental paradigm applied in all subsequent experiments, which focusses on argument linking processes in the *nominal*, rather than the verbal domain, and makes use of deverbal nominalisations as predicates, as already pointed out in the Introduction. Section 2.1 summarises relevant general issues related to parallels and differences in the argument linking processes in the verbal and nominal domain and section 2.2 focusses on crucial argument linking phenomena arising with German *-ung*-nominalisations of syntactically transitive verbs and postnominal genitives. Section 2.3 presents the semantic account of these patterns developed by Ehrich and Rapp (2000), in which the lexical semantic notion of change of state/affectedness plays a prominent role; section 2.4 reinterprets Ehrich and Rapp's account in terms of semantic prominence of the direct object argument as defined by the Affectedness Hierarchy of Beavers (2010, 2011). After a discussion of the syntactic status of genitive arguments of German *-ung*-nominalisations in section 2.5, section 2.6 outlines the rationale for taking a detour via deverbal nominalisations for the investigation of verbal argument linking processes. Section 2.7 presents and discusses an acceptability judgement experiment intended to validate the experimental paradigm and empirically test the predictions derived from the reinterpretation of Ehrich and Rapp's semantics-based theory based upon Beavers' Affectedness Hierarchy.

## 2.1 Argument linking: from verbs to nouns

Among the major building blocks of language are the major parts of speech, i.e. verbs, nouns, adjectives and prepositions (V, N, A and P) – a definition of the properties which distinguish these major lexical categories from one another on the one hand, and those features which some of them share in common on the other hand is central to a theory of grammar. Chomsky (1970) uses the two features  $[\pm N]$  ('nominal') and  $[\pm V]$  ('verbal') to classify these categories as  $[+V, -N]$  for verbs,  $[-V, +N]$  for nouns,  $[+V, +N]$  for adjectives and  $[-V, -N]$  for prepositions<sup>1</sup>. Wunderlich (1996) points out that this feature decomposition lacks a conceptual basis and does not characterise some of the most relevant relations between the categories correctly. His alternative classification system makes use of the two features  $[\pm dep]$  for 'referentially dependent' and  $[\pm art]$  for 'articulated'; nouns and verbs share the feature specification  $[-dep]$ , since items from both classes can be anchored in extra-linguistic contexts via referential arguments, making them referentially independent. Adjectives as well as prepositions are referentially dependent ( $[+dep]$ ) and need support by the linguistic context. In Wunderlich's system, nouns and verbs are distinguished by the specification for the  $[art]$  feature: verbs and prepositions are defined as  $[+art]$  because they are specified for (multiple) obligatory arguments, while nouns and adjectives lack such obligatory arguments and are thus specified as  $[-art]$ .

On this view, the main difference between verbs and nouns lies in the distinction between rich argument structure underlying verbs and impoverished argument structure associated with nouns. Stiebels (2006, pp. 170–171) illustrates the difference in argument structure possibilities between the two word classes in the predicate-argument structure notation of *Lexical Decomposition Grammar* (e.g. Wunderlich, 1997): the decompositions in (66) exemplify different verbs with increasingly complex argument structure, where the argument  $s$  represents the event/situational variable, i.e. the verb's referential argument. The argument structure is derived from the lexical decompositions by means of  $\lambda$ -abstraction<sup>2</sup> of the variables present in the decompositional structures, placing the lowest-ranked (i.e. most embedded) argument at the leftmost position and the highest argument (the referential argument) at the rightmost. Note that the order of the  $\lambda$ -abstractors matters for argument realisation, as discussed in the following section. As the examples in (66) show, verbs may show any number of structural arguments from none (66a) to three (66d), in some languages even up to five (Stiebels, 2006, p. 170).

(66) a. rain:  $\lambda s \text{ RAIN}(s)$

<sup>1</sup>'+' and '-' are to be understood as indicating whether the respective feature holds or does not hold for a given word class, respectively.

<sup>2</sup>The 'Lambda calculus' is a formal and universal system for the representation of functions introduced by Church (1936) (see Alama, 2015). It is widely used in a number of fields and can be applied to represent the meaning of expressions of artificial and natural languages in an abstract and compact way. As a first example, take the simple mathematical expression  $x - 2$ , where  $x$  is a variable; in  $\lambda$ -notation, this can be represented as  $\lambda x [x - 2]$ , where the  $\lambda$  binds the variable  $x$ . To assign a value to  $x$ , we can write  $\lambda x [x - 2] (5)$ , which indicates that the value '5' is to be substituted for any occurrences of  $x$  in the expression  $x - 2$ ; this yields  $5 - 2 = 3$ . Abbott (2003, p. 79) gives an example for its use in linguistic semantics, noting that the "function denoted by  $\lambda x [\text{Mary loves } x]$  maps individuals whom Mary loves to the value True, and others to the value False".

- b. cry:  $\lambda x \lambda s \text{CRY}(x)(s)$
- c. read:  $\lambda y \lambda x \lambda s \text{READ}(x, y)(s)$
- d. give:  $\lambda z \lambda y \lambda x \lambda s [\text{ACT}(x) \ \& \ \text{BECOME}(\text{POSS}(y, z))](s)$

On the other hand, a prototypical noun – i.e. a non-relational noun denoting some (more or less) concrete entity – such as *cat* denotes a one-place predicate as illustrated in (67a). Some nouns such as kinship terms (67b) or body part expressions (67c) are inherently relational and thus denote two-place predicates (see Stiebels, 2006, p. 171).

- (67) a. cat:  $\lambda u \text{CAT}(u)$   
 b. mother:  $\lambda v \lambda u \text{MOTHER}(u, v)$   
 c. nose:  $\lambda v \lambda u [\text{NOSE}(u) \ \& \ \text{PART\_OF}(u, v)]$

Stiebels uses the metaphor of ‘being in rags’ in referring to this impoverished argument structure associated with prototypical nouns – however, she stresses and illustrates that one can find nominal parallels to operations on verbal arguments structure (e.g. voice/diathesis operations such as causative or applicative) and that in some instances nouns can actually ‘be in riches’ in exhibiting more complex argument structure, just as verbs may. Investigating nouns with more complex argument structure involves moving away from prototypical nouns; as Wunderlich (1996, p. 18) remarks, “[n]ouns exhibit the greatest sortal variation in that they may refer to any spatial or temporal or even more abstract entity” as exemplified by basic and underived nouns such as *war*, *event*, *game* or *thunderstorm*.

The full complexity of verbal argument/event structure, however, only comes to bear in nouns which are derived from verbs, as in deverbal nominalisations like *destruction*, *assassination* or *examination*, which are morphologically derived from the respective base verbs *destroy*, *assassinate* and *examine* (see Alexiadou, 2010, for an introduction to issues related to nominalisations). Grimshaw (1990) argues that such deverbal nominalisations may inherit the full argument and event structure of the corresponding base verbs, as in (68).

- (68) a. The instructor’s examination of the papers took a long time.  
 b. The enemy’s destruction of the city was awful to watch.

Grimshaw further shows that deverbal nominalisations can be systematically ambiguous between what she calls ‘complex event nominals’ (CENs) on the one hand and ‘simple event nominals’ (SENs) and ‘result nominals’ (RNs) on the other hand – while the former are associated with true verbal event and argument structure, the latter actually lack these, though they may still be interpreted as events or actions. When derived nominals are interpreted as SENs/RNs, these are internally structured and behave like basic, underived nouns without verbal event/argument structure, as in (69).

- (69) The examination/exam is on the table.

The presence of the base verb's aspectual structure can be witnessed by the compatibility of CENs with the adjectival equivalents of event-modifying adverbs such as *frequently* or *constant*: in (70b), *frequent/constant* can modify the aspectual structure inherited from the base verb *examine* and the temporal PPs *for an hour* and *in an hour* are both compatible with the construction in (70b), due to the possible interpretation of *examine patients* as either telic or atelic (see Grimshaw, 2011, p. 1300).

- (70) a. They frequently/constantly examine patients in/for an afternoon.  
 b. The frequent/constant examination of patients in/for an afternoon leads to better diagnoses.

On Grimshaw's view, the presence of aspectual structure in CENs can also go hand in hand with the inheritance of the verb's argument structure and thus with the obligatory realisation of arguments: since the base verb *examine* takes an obligatory direct object (cf. (71a)), this must also be present with the CEN *examination*, explaining the unacceptability of (71b) (from Grimshaw, 2011, p. 1301).

- (71) a. \* They frequently/constantly examine in/for an afternoon.  
 b. \* The frequent/constant examination in/for an afternoon leads to more accurate diagnoses.

Thus, deverbal event nominalisations are a kind of hybrid category, incorporating purely nominal features (e.g. they can function as heads of NP arguments of verbs and are modified by adjectives) as well as verbal properties by inheriting parts of the lexical verbal event and argument structure, with different types of deverbal nominal constructions exhibiting varying degrees of 'verbiness' (see Ehrich, 2002, for discussion). When it comes to linking the arguments associated with the verbal lexical structure within an NP headed by a deverbal event nominalisation, grammar is facing a challenge though: since the nominal linking resources are based upon the 'rags' associated with underived nouns, accommodating the verbal 'riches' is not always a trivial enterprise. Since basic nouns are at most associated with two arguments (see (67) above), the nominal linking inventory is often under-equipped to deal with complex verbal argument structure, as discussed by Stiebels (2006, p. 172):

"Since the referential argument does not play a role for argument linking in non-predicative uses of nouns, structural linking is generally rather limited, often restricted to one case/agreement linker, which, in addition, is barred from being doubled in many languages. These restrictions on structural linking also have an impact on the realization of inherited arguments in nominalizations. Often, only one argument may receive a structural nominal linker."

In this situation, the grammar must sometimes provide ways to compensate for this lack of structural linkers in nouns and sometimes the different arguments of a verb compete for one and the same argument position. In the following sections, I will outline how such a competition in German deverbal event nominalisations can be used to investigate aspects of the role of affectedness on argument linking with nominalisations.

## 2.2 Linking patterns in German deverbal *-ung* nominalisations

In (Modern High) German, two positions are available for the realisation of genitive possessors and arguments within an NP: the prenominal position illustrated in (72a) and the postnominal one exemplified in (72b). These two positions are available in NPs headed by either underived nouns such as *Krone/Tochter* (*crown/daughter*), whether relational or not, as well as in NPs headed by nominalisations, such as *Ermordung* (*assassination*).

- (72) a. Cäsar-s      Krone/Tochter/Ermordung  
           Caesar-GEN crown/daughter/assassination  
           ‘Caesar’s crown/daughter/assassination’  
       b. die Krone/Tochter/Ermordung      des      Konsul-s  
           the crown/daughter/assassination the.GEN consul-GEN  
           ‘the crown/daughter/assassination of the consul’

Importantly, prenominal genitives are limited to proper names as in (72a); the realisation of full genitive NPs such as *des Konsuls* (*the consul’s*) as in (73a) is usually judged as highly marked or ungrammatical in modern German. As shown by (72b) and (73b), the postnominal genitive position is fully productive since it can be filled by any kind of genitive NP.

- (73) a. ?\* des      Konsul-s      Krone/Tochter/Ermordung  
           the.GEN consul-GEN crown/daughter/assassination  
           ‘the consul’s crown/daughter/assassination’  
       b. die Krone/Tochter/Ermordung      Cäsar-s  
           the crown/daughter/assassination Caesar-GEN  
           ‘the crown/daughter/assassination of Caesar’

The examples in (74) illustrate a number of additional genitive realisation options and restrictions: where the relational nature of the head noun licenses two genitives, these can be realised simultaneously with one genitive (proper name) occurring in pre- and the other one in postnominal position as in (74a). With deverbal event nominalisations of syntactically transitive verbs (such as *Zerstörung/destruction*), the much more natural way of realising the argument linked as the subject with the base verb is within a PP headed by *durch* (*by*), which follows the postnominal genitive (74b). Note that postnominal genitives must be strictly adjacent to the head noun as exemplified by (74c) and that doubling of genitives in either position is not possible (74d)<sup>3</sup>.

- (74) a. Cäsar-s      Zerstörung der      Stadt  
           Caesar-GEN destruction the.GEN city.GEN  
           ‘Caesar’s destruction of the city’

<sup>3</sup>Hartmann and Zimmermann (2003), however, argue that postnominal genitives may be doubled under specific conditions, especially when contrastive focus is involved.

- b. die Zerstörung der Stadt durch Cäsar  
the destruction the.GEN city.GEN by Caesar.ACC  
'the destruction of the city by Caesar'
- c. \* die Zerstörung durch Cäsar der Stadt  
the destruction by Caesar.ACC the.GEN city.GEN  
'the destruction by Caesar of the city'
- d. \* die Zerstörung der Stadt Cäsars  
the destruction the.GEN city.GEN Caesar-GEN  
'the destruction of the city of Caesar'

These patterns raise a number of questions about the syntactic structure and mechanisms underlying the German NP, including the syntactic status of prenominal genitives (i.e. are they best treated as syntactic heads or full phrases?) and the derivation of pre- and post-nominal genitives (e.g. are they base-generated in their respective positions or moved there from other position?). Though these are central issues related to the nature of German NPs, a discussion of these issues would go far beyond the scope of the current work and is not mandatory for my current purposes – thus, I will largely abstract away from the syntactic details underlying the relevant constructions. The syntactic properties of the German NP and genitive constructions in particular have been discussed in a number of studies, including Haider (1988), Olsen (1991), De Wit and Schoorlemmer (1996), Lindauer (1998), Hartmann and Zimmermann (2003), Bücking (2011), Fuß (2011), among others.

Also, I will restrict the subsequent overview of genitive linking to the postnominal genitive position and thus also to the realisation of a single genitive phrase, since this is the only relevant case for the experimental investigations which are to follow. One of the fundamental issues with argument linking in nominals is how the linking of arguments of polyadic base verbs is regulated; as discussed in the previous section, the impoverished linking resources available with nouns often do not suffice to accommodate all arguments of a verb via structural linkers such as genitive case. Considering that the prenominal genitive position within the German NP domain is subject to the severe restrictions discussed above, the question of which of the arguments of a syntactically transitive verb<sup>4</sup> can be realised as the sole postnominal genitive arises.

At this point, a brief terminological aside may be in order: in the following, I will not use thematic role labels such as 'Agent' or 'Patient' in referring to genitive phrases representing arguments of a deverbal nominalisation's base verb. As the discussion in chapter 1 showed, there is no one-to-one correspondence between such categories and grammatical functions and they are often used as default labels for arguments which escape easy semantic classification and do not match the respective role label's concept at all. Rather, I will use 'GenS' in referring to genitives representing the argument linked as subject of the underlying base verb (*genitivus subjectivus*) and 'GenO' for genitives representing its direct object (*genitivus objectivus*), and I will make use of Proto-role labels for associated thematic roles. Also,

<sup>4</sup> I am focussing on syntactically transitive verbs here and not discussing linking patterns of verbs of higher valency, such as ditransitive verbs.

since there are other types of deverbal nominalisations (e.g. participant nominals such as *writer*, *employer* or *employee* – for an overview see Grimshaw, 2011), I will henceforth use the term ‘eventive (deverbal) nominalisations’ to refer to the nominalisation type of interest in the current study, where ‘eventive’ is intended to comprise all kinds of aspectual verb classes (e.g. states, processes, accomplishments and achievements).

On traditional morphosyntactic accounts, the thematic role or the lexical semantic properties do not play a role in determining the linking of a base verb’s arguments as genitives with eventive nominalisations; rather, it is assumed that an eventive nominalisation inherits the base verb’s argument which carries structural accusative case and realises this argument with structural genitive (for discussion, see Ehrich & Rapp, 2000, section 3.1). Such an approach can explain the pattern illustrated in example (75): here, *des Generals* can ONLY be interpreted as GenO, i.e. as the base verb’s direct object – as the one being dismissed/detained/expatriated/assassinated (i.e. Proto-Patient). An interpretation as GenS (Proto-Agent) is not available.

- (75) die Absetzung/Verhaftung/Verbannung/Ermordung des General-S\*GenS/GenO  
 the dismissal/detention/expatriation/assassination the.GEN general-GEN  
 ‘the dismissal/detention/expatriation/assassination of the general’

Example (76) illustrates this linking pattern within a discourse context: (76a) introduces the situation and unambiguously links the *König/king* as the subject and the *General/general* as the direct object (corresponding to the roles of Proto-Agent and Proto-Patient, respectively). Assuming that (76b) and (76c) are both intended to be coherent continuations of this discourse situation, (76b) is consistent with (76a), since it links the argument there realised as direct object as GenO; (76c), on the other, represents an inconsistent continuation, since the intended interpretation of the postnominal genitive is GenS, which is not available. The *king* can only be understood to represent the Proto-Patient participant underlying GenO, so the *king’s* interpretation is here incompatible with the established discourse context.

- (76) a. *Target situation*:

Der König hat den General  
 the.NOM king.NOM has the.ACC general.ACC  
 abgesetzt/festgenommen/verbannt/ermordet.  
 dismissed/detained/expatriated/assassinated

‘The king dismissed/detained/expatriated/assassinated the general.’

- b. a. = die Absetzung/Verhaftung/Verbannung/Ermordung des General-s<sub>GenO</sub>  
 the dismissal/detention/expatriation/assassination the.GEN general-GEN  
 ‘the dismissal/detention/expatriation/assassination of the general’  
 c. a. ≠ die Absetzung/Verhaftung/Verbannung/Ermordung des König-s\*<sub>GenS</sub>  
 the dismissal/detention/expatriation/assassination the.GEN king-GEN  
 ‘the dismissal/detention/expatriation/assassination of the king’

While a morphosyntactic inheritance account can explain this asymmetric linking pattern, it fails to account for the fact that this asymmetry is not found with all syntactically transitive verbs. In contrast to *des Generals/the general's* in (75), in (77) *des Polizisten/the policeman's* can represent either GenO or GenS, with GenO generally being the preferred interpretation.

- (77) die Befragung/Demütigung/Ermahnung/Unterstützung/Bewunderung des  
 the interrogation/humiliation/admonition/support/admiration the.GEN  
 Polizist-en<sub>GenS/GenO</sub>  
 policeman-GEN  
 'the interrogation/detention/admonition/support/admiration of the policeman'

(78a) sets up a discourse context parallel to (76a), with *der Polizist/the policeman* as subject and *den Fahrer/the driver* as accusative-marked direct object of *interrogate/humiliate/admonish/support*. While these verbs are equally syntactically transitive as the verbs in (76), both (78b) with GenO and (78c) with GenS are consistent continuations of (78).

- (78) a. *Target situation:*

Der Polizist hat den Fahrer  
 the.NOM policeman.NOM has the.ACC driver.ACC  
 befragt/gedemütigt/ermahnt/unterstützt/bewundert.  
 interrogated/humiliated/admonished/supported/admired

'The policeman interrogated/humiliated/admonished/supported/admired the driver.'

- b. a.= die Befragung/Demütigung/Ermahnung/Unterstützung/ Bewunderung  
 the interrogation/humiliation/admonition/support/ admiration  
 des Fahrer-s<sub>GenO</sub>  
 the.GEN driver-GEN  
 'the interrogation/humiliation/admonition/support/ admiration of the driver'
- c. a.= die Befragung/Demütigung/Ermahnung/Unterstützung/ Bewunderung  
 the interrogation/humiliation/admonition/support/ admiration  
 des Polizist-en<sub>GenS</sub>  
 the.GEN policeman-GEN  
 'the interrogation/humiliation/admonition/support/admiration of the policeman'

The failure of purely morphosyntactic accounts to capture this difference in linking options between different verbs suggests that the linking mechanisms with such eventive nominalisations is influenced by semantic factors. The next section introduces the semantic account presented by Ehrich and Rapp (2000).



## 2.3 Change of state as a semantic determinant of argument linking in eventive *-ung* nominalisations

Ehrich and Rapp (2000) investigate the semantic aspects determining the linking possibilities of genitive arguments in German deverbal eventive nominalisation formed with *-ung*. In German, *-ung* is the most frequent and a fully productive derivational suffix for forming eventive nominalisations from verbs; it yields feminine nouns and is semantically largely transparent. Importantly, Ehrich and Rapp argue that *-ung* nominalisations inherit the full lexical-semantic structure of their base verb. Making use of lexical decompositions and a notation similar to the one introduced in section 1.4.1, (79) illustrates the decompositions for *-ung* nominalisations based upon verbs of different classes. (79a) and (79b) are the representations of eventive nominalisations of dyadic state and activity verbs (Ehrich & Rapp, 2000, p. 281), with each verb class having a specific referential argument (*s* for states, *r* for activities and *e* for change of state predicates as in (79c)). Note that the decomposition of the *lexical-semantic structure* (LSS) is enclosed within square brackets and the variables appearing in these are again made available for the syntax via  $\lambda$ -abstraction within the *argument structure* part: the left-to-right order within the argument structure reflects the depth of embedding within the LSS (from most deeply embedded/lowest ranked argument to the highest ranked argument) and the parentheses around the *x* and *y* arguments indicate that their syntactic realisation is optional. In Ehrich and Rapp's approach, the corresponding base verbs of (79a) and (79b) have an identical LSS, the only difference lying in the argument structure: *x* and *y* are optional with eventive *-ung* nominalisations, but are obligatory (and thus not parenthesised) in the argument structure of verbs.

- (79) a. *Bewunderung*:  $(\lambda y) (\lambda x) \lambda s [POSS ((x, y) s)]$   
 b. *Verfolgung*:  $(\lambda y) (\lambda x) \lambda r [DO ((x, y) r)]$   
 c. *Herstellung*:  $(\lambda y) \lambda r [DO ((x, y) r) \& BEC ((BE ((y) s)) e)]$   
 d. *Ermordung*:  $(\lambda y) \lambda r [DO ((x, y) r) \& BEC ((NOT (BE ((y) s))) e)]$

A look back at the examples in (77) and (78) reveals that all of the respective base verbs are either stative (such as the psychological predicate *bewundern* (*admire*)) or activity verbs (*be-fragen/demütigen/ermahnen/unterstützen* (*interrogate/humiliate/admonish/support*)). According to the decompositions in (79a) and (79b), the respective argument structures feature the GenS argument *x as well as* the GenO argument *y*: while none of these have to be realised (due to their optional syntactic status), either of these *can* thus be realised as a (postnominal) genitive, since both are made available for syntactic realisation via  $\lambda$ -abstraction.

Another look at the examples in (75) and (76), which only allow for GenO linking, shows that all of the respective verbs (*absetzen/festnehmen/verbannen/ermorden* (*dismiss/detain/expatriate/assassinate*)) involve a change of state component, which is encoded by a *BECOME* operator in the decompositions of eventive nominalisation of causative verbs of creation such as *Herstellung/Fertigstellung* (*creation/completion*) in (79c) or destruction such as *Zerstörung/Ermordung* (*destruction/assassination*) in (79d) (Ehrich & Rapp, 2000, p. 284). While

Ehrich and Rapp assume identical semantic decompositions of verbs and their eventive *-ung* nominalisations, they postulate separate linking rules for each word class and it is the presence of *BECOME* in the semantic decomposition which makes the difference between those nominalisations which allow for the linking of GenO as well as of GenS and those which only license linking of GenO. The crucial linking rule explaining this difference says that if the semantic decomposition does NOT contain a *BECOME* operator (i.e. a change of state component), *all* of the thematic arguments appear in the respective argument structure (as in (79a) and (79b)), whereas *only the lowest-ranked thematic argument is made available for syntactic linking if the LSS includes such a BECOME element* (see Ehrich & Rapp, 2000, p. 276). Since the lowest-ranked argument corresponds to the most deeply embedded argument in the semantic decomposition structure, only the *y* argument corresponding to the verb's direct object is available for linking as GenO in the representations in (79c) and (79d). Since the *x* argument is lacking in the nominalisations' argument structure, linking of GenS is not an option, rendering the genitive unambiguous.

Note that Stiebels (2006) explains the priority of the lowest argument for linking in nominals by their (cross-linguistic) preference for strict successive argument saturation, starting bottom-up with the lowest-ranked of all arguments. She notes that while skipping of arguments is strongly dispreferred, lexical-semantic/aspectual properties of base verbs such as identified by Ehrich and Rapp can license a violation of this strictly successive bottom-up linking algorithm.

Two central assumptions of Ehrich and Rapp (2000) are a) that the postnominal genitive position is a structural argument position within NP (they actually assume that it is the only one – see p. 275) and b) that GenO as well as GenS represent the respective arguments of the underlying base verb, rather than an argument-like adjunct. The argument status of GenO/GenS holds despite the fact that their syntactic realisation is optional according to the argument structure of eventive *-ung* nominalisations. The attribution of a full argument status to postnominal GenO and/or GenS is not uncontroversial within the literature on German eventive nominalisations, and I will briefly comment on some of the issues in section 2.5 below.

As Ehrich and Rapp point out, each of the basic predicates used in the decomposition structures (such as *DO* (*x, y*) or *POSS* (*x, y*)) opens up a 'local' thematic hierarchy (see also the discussion in section 1.3.2 of the previous chapter) with the first argument outranking the second one. These hierarchies are referred to by the linking rules. The difference between verbal and nominal linking lies in the focus of the respective linking rules: the verbal linking rules always prioritise the thematic structure opened up by these basic predicates, whereas the causal structure does not play a role. As a consequence, the presence of a *BECOME* operator does not influence the argument linking with verbs. Where linking conflicts arise in the verbal domain (such as with three-place verbs), the arguments of the *DO* component are always prioritised over arguments of more deeply embedded predicates. Interestingly, the nominal linking pattern observed with eventive *-ung* nominalisations is diametrically opposed to the verbal one: here, for predicates involving a change of state and thus a *BECOME* operator, the priority is given to the resultant state part of the LSS represented by *BE*. Ehrich

and Rapp (2000, p. 300) conjecture of the origin of this systematic difference between verbal and nominal linking (my translation, emphasis in the original):

“[t]he verb – as a prototypically dynamic category – gives preference to the *activity* part of the LSS. In contrast, a deverbal nominalisation – which represents the transition from the verb to the prototypically static category *noun* – prioritises the result state component of the LSS.”

Finally, I will briefly comment upon the relation between the observed influence of a change of state component with the LSS of an eventive *-ung* nominalisation and conceptual/semantic properties inherent to the head noun of the postnominal genitive NP, such as animacy. Of course such features can give soft pragmatic or hard semantic cues for the interpretation of genitives linked to eventive nominalisations when these appear in isolation, i.e. without any context determining the genitive’s interpretation. With a nominalisation like *Behandlung/treatment*, which does not entail a necessary change of state and thus allows for linking of GenO and GenS, *des Arztes* in *die Behandlung des Arztes* (*the treatment of the doctor*) may preferentially be interpreted as GenS on pragmatic grounds, though nothing speaks against understanding *the doctor* as GenO; after all, doctors need to be treated themselves from time to time. In *die Behandlung des Bruches* (*the treatment of the fracture*), the interpretation of *Bruch/fracture* is trivially restricted to GenO, since inanimate entities can hardly treat someone else.

The following two examples show that the differential genitive linking patterns found with nominalisations based on verbs without (80) and with a change of state component (81) hold independently of the animacy status of the base verb’s direct object:

(80) a. *Target situation:*

Der Doktor behandelte den Patienten / den Bruch.  
the.NOM doctor.NOM treated the.ACC patient.ACC / the.ACC fracture.ACC

‘The doctor treated the patient / fracture.’

b. a. = die Behandlung des Patient-en<sub>GenO</sub> / des Bruch-s<sub>GenO</sub>  
the treatment the.GEN patient-GEN / the.GEN fracture-GEN

‘the treatment of the fracture’

c. a. = die Behandlung des Doktor-s<sub>GenS</sub>  
the treatment the.GEN doctor-GEN

‘the treatment of the doctor’

(81) a. *Target situation:*

Der Manager beseitigte den Konkurrenten / das Problem.  
the.NOM manager.NOM removed the.ACC competitor.ACC / the.ACC problem.ACC

‘The manager removed the competitor / problem.’

- b. a.= die Beseitigung des Konkurrent-en<sub>GenO</sub> / des Problem-s<sub>GenO</sub>  
 the removal the.GEN competitor-GEN / the.GEN problem-GEN  
 'the removal of the competitor / problem'
- c. a.≠ die Beseitigung des Manager-s\*<sub>GenS</sub>  
 the removal the.GEN manager-GEN  
 'the removal of the manager'

This may not come as a surprise, however, and a more conclusive proof of the independence of the observed LSS-related linking restrictions would have to show that the animacy status of GenS does not allow to override these. (82) suggests that the two factors indeed do not interact in this way: here, the subject *the rock* of the underlying change of state base verb *kill* is inanimate, but it still cannot be interpreted as GenS as indicated by (82b). If the animacy of the arguments had a critical influence on the observed linking restrictions with eventive nominalisations of change of state verbs, it should probably surface in this constellation; since it is clear that an inanimate entity like a rock cannot be killed, an interaction of animacy and LSS-related factors may save (82b) and render it licit by means of a kind of last resort operation which overrides the restricting nominal linking rule and shifts the interpretation of *the rock* to GenS. This is not the case, however, and the construction in (82b) is in fact conceptually anomalous, exactly because a rock cannot be killed. This underscores the primacy of the semantic linking factors induced by the verbal LSS over argument-inherent semantic/conceptual properties.

(82) a. *Target situation:*

Der Stein tötete den Bergsteiger.  
 the.NOM stone.NOM killed the.ACC climber.ACC

'The rock killed the climber.'

- b. a.≠ (#) die Tötung des Stein-s\*<sub>GenS</sub>  
 the killing the.GEN rock-GEN  
 'the killing of the stone'

## 2.4 Determining linking in *-ung* nominalisations: a case of semantic prominence induced by degrees of affectedness?

Ehrich and Rapp (2000, pp. 285 ff.) discuss a linking pattern which – on first sight – seems to run counter to their semantics-based theory of argument linking with eventive *-ung* nominalisations. The critical examples provided by them all involve nominalisations based on verbs of modification such as *Kürzung/Modifizierung/Renovierung* (*abridgement/modification/renovation*); while these all clearly imply a change of state for their direct object, Ehrich and Rapp note that a GenS interpretation "is not completely unavailable" (p. 285). Thus, *der*

*Redakteur/the editor* in (83a) and *der Architekt/the architect* in (83b) can be linked as GenS despite of the change clearly implied for the verb's objects.<sup>5</sup>

- (83) a. die Kürzung / Modifizierung des Artikel-s<sub>GenO</sub> / ✓?des  
 the abridgement / modification the.GEN article-GEN / the.GEN  
 Redakteur-s<sub>GenS</sub>  
 editor-GEN  
 'the abridgement/modification of the article/the editor'
- b. die Renovierung / Umgestaltung des Gebäude-s<sub>GenO</sub> / ✓?des  
 the renovation / remodelling the.GEN building-GEN / the.GEN  
 Architekt-en<sub>GenS</sub>  
 architect-GEN  
 'the renovation/remodelling of the building/the architect'

The authors argue that these examples actually do not pose a problem for their approach and identify the aspectual ambiguity of the respective base verbs as the source licensing the GenS interpretation in these cases. A verb like *kürzen* can either be interpreted as atelic, focussing on the process component of the verb's semantic structure, or as telic, in which case the result state component is made salient. In (84a), the atelic activity reading is forced by the time span adverbial *for three hours* and in (84b), the telic accomplishment reading is obtained by the time frame modifier *in three hours* (examples are from Ehrich & Rapp, 2000, p. 286). Since such verbs are associated with two LSS and *-ung* nominalisations inherit the base verb's LSS, Ehrich and Rapp explain the possibility of linking GenS with such nominals by their dual character: the activity variant lacks a *BECOME* operator and is thus compatible with a GenS interpretation.

- (84) a. *Atelic (activity) interpretation:*  
 Der Redakteur kürzte den Artikel drei Stunden lang.  
 the.NOM editor.NOM abridged the.ACC article.ACC three hours long  
 'The editor abridged the article for three hours.'
- b. *Telic (accomplishment) interpretation:*  
 Der Redakteur kürzte den Artikel in nur einer Stunde.  
 the.NOM editor.NOM abridged the.ACC article.ACC in only one hour  
 'The editor abridged the article in one hour only.'

In my opinion, a related interpretation of these facts can be obtained by explaining them within the scalar approach to degrees of affectedness of Beavers (2010, 2011) introduced in section 1.5.1: verbs such as *abridge/modify/renovate/remodel* belong to the class of degree achievements like *cool* (Hay et al., 1999; Kennedy & Levin, 2008), which can be interpreted as referring to a quantised change with a concrete goal state on the associated scale as in

<sup>5</sup>Note that I am using the tick mark (✓) here as a means of explicitly indicating acceptability of a construction to allow for a more graded representation of (my personal) intuitions. See also the comments on example (86) below.

(85a) or to a non-quantised change where the exact result state on the scale is left open as in (85b) – on the first reading the expression is telic, whereas it is atelic on the second interpretation. Similarly, all of the base verbs of the eventive nominalisations in (83) can imply quantised or non-quantised changes; after all, something can be abridged, modified, renovated and remodelled to a very specific degree, or to some non-specified degree.

- (85) a. Before serving, the cook cooled the soup to 65°C in/?for five minutes.  
 b. Before serving, the cook cooled the soup ?in/for five minutes.

With the base verbs in (83) reinterpreted as non-quantised change predicates, we can now actually map the linking patterns found with eventive *-ung* nominalisations onto Beavers's Affectedness Hierarchy as stated in (49), which is repeated below. Those eventive nominalisations which do not license linking of GenS as in (75) all entail a quantised change for their direct object and thus the theme  $x$  transitions to a specific state  $g_\phi$  on the associated scale as indicated in (49a). The nominalisations of degree achievement verbs in (83) which allow for a (dispreferred) GenS as well as a GenO interpretation of the postnominal genitive correspond to non-quantised change verbs as in (49b), i.e. they are associated with an unspecified result state  $g$  on the respective scale. Further, the base verbs of those nominalisations in (77), which license linking of GenS and GenO and do not imply any necessary change of state, correspond to the potential change verbs (e.g. surface contact/impact verbs) in (49c) and verbs unspecified for change (e.g. states and certain activity verbs) in (49d). The former are associated with a latent scale argument  $s$  and their objects qualify as force recipients and may thus potentially undergo a change of state, but neither of these two verb classes entails a necessary change.

- (49) a.  $x$  undergoes a **quantized** change iff  $\phi \rightarrow \exists e \exists s [\text{result}'(x, s, g_\phi, e)]$   
 (e.g. accomplishments/achievements: *break, shatter, destroy, devour x*)  
 b.  $x$  undergoes a **non-quantized** change iff  $\phi \rightarrow \exists e \exists s \exists g [\text{result}'(x, s, g, e)]$   
 (e.g. degree achievements: *widen, cool, lengthen, cut, slice x*)  
 c.  $x$  has **potential** for change iff  $\phi \rightarrow \exists e \exists s \exists \theta [\theta(x, s, e)]$   
 (e.g. surface contact/impact: *wipe, scrub, rub, punch, hit, kick, slap x*)  
 d.  $x$  is **unspecified** for change iff  $\phi \rightarrow \exists e \exists \theta' [\theta'(x, e)]$   
 (e.g. other activities/states: *see, laugh at, smell, follow, ponder, ogle x*)

The examples in (86) illustrate the possible genitive linking patterns with eventive *-ung* nominalisations for each of the four levels of the Affectedness Hierarchy, where *des Arztes/of the doctor* is always intended to represent GenS and *des Patienten/of the patient* GenO. Note that I am providing my own graded intuitions about the ease with which each of the interpretations can be obtained in these examples, with two checkmarks (✓✓) indicating straightforward availability. Starting with the lowest level exemplified in (86a), nominalisations of base verbs which are unspecified for change allow equal linking of GenO and GenS. With base verbs which have potential for change as in (86b), both of these two options are available as well, but – in my intuition – less preferred than GenO (hence only one checkmark). With

non-quantised change base verbs as in (86c), the GenS interpretation is still possible, but may be less readily available as with the two lower levels. At the highest level with base verbs indicating quantised change, linking of GenS becomes impossible (86d). Note that the base verbs in these examples come from different conceptual fields and include mental verbs (*bewundern/ermutigen* – *admire/encourage*), verbs of change of location (*einliefern* – *hospitalise*) and verbs of physical manipulation and impact such as *behandeln/untersuchen/sedieren* (*treat/examine/sedate*). The fact that the linking possibilities are independent of the real-world aspects associated with the change (not) entailed by the verb is actually predicted by Beavers's theory, since it abstracts away from the type of scale and focusses on its structure<sup>6</sup>.

(86) a. *Unspecified for change:*

die Bewunderung / Begrüßung ✓✓des Patient-en<sub>GenO</sub> / ✓✓des Arzt-es<sub>GenS</sub>  
 the admiration / greeting the.GEN patient-GEN / the.GEN doctor-GEN  
 'the admiration/greeting of the patient/the doctor'

b. *Potential for change:*

die Behandlung / Untersuchung ✓✓des Patient-en<sub>GenO</sub> / ✓des Arzt-es<sub>GenS</sub>  
 the treatment / examination the.GEN patient-GEN / the.GEN doctor-GEN  
 'the treatment/examination of the patient/the doctor'

c. *Non-quantised change:*

die Sedierung / Ermutigung ✓✓des Patient-en<sub>GenO</sub> / ✓?des Arzt-es<sub>GenS</sub>  
 the sedation / encouragement the.GEN patient-GEN / the.GEN doctor-GEN  
 'the sedation/encouragement of the patient/the doctor'

d. *Quantised change:*

die Heilung / Einlieferung ✓✓des Patient-en<sub>GenO</sub> / \*des Arzt-es<sub>GenS</sub>  
 the cure / hospitalisation the.GEN patient-GEN / the.GEN doctor-GEN  
 'the cure/hospitalisation of the patient/the doctor'

The Affectedness Hierarchy seems to provide a natural alternative approach to the argument linking patterns found with *-ung* nominalisations and while it offers a somewhat different perspective on Ehrich and Rapp's theory, I believe the two approaches are quite compatible; remember that Beavers's scalar model ultimately defines affectedness in terms of the strength of truth conditions about the effect a verb entails for a participant, with more specific effects associated with (monotonically) stronger truth conditions. Stronger truth conditions and thus higher degrees of affectedness, in turn, can be interpreted as *increasing semantic prominence* of the participant realised as the object, with consequences for argument linking as discussed by Beavers (2010).

<sup>6</sup>While Beavers discusses a number of different types of change (property, existence and location), he does not address change in possession, status and mental state, but assumes that (some of) the latter can be modelled equivalently (see Beavers, 2011, p. 366 and his footnote 1 on page 339).

Thus, Beavers' scalar model of affectedness provides immediate links to Ehrich and Rapp's theory of nominal linking: the categorical influence of the *BECOME* operator can be reinterpreted as degree of affectedness on the Affectedness Hierarchy, potentially allowing to explain more graded judgements, if the intuitions in (86) are representative – the more semantically prominent GenO is with respect to affectedness, the less acceptable is the GenS interpretation of a genitive. In addition, the fact that the priority of GenO rises with increasing semantic prominence fits nicely with Ehrich and Rapp's conjecture that in contrast to verbs, nouns prioritise the result state component of the underlying verb's LSS, explaining why the relevant semantic prominence is computed with respect to GenO, rather than GenS.

Note that another complementary perspective upon the relation between affectedness and argument linking with eventive nominalisations is provided by the notion of semantic transitivity as discussed in section 1.5.2.1 in the previous chapter. Since affectedness of the object and semantic transitivity are closely related, Beavers (2011, p. 362) points out that his hierarchy can be interpreted as arranging verbs in the order of their degree of transitivity. In the 'bipolar' approaches to semantic transitivity of Blume (2000) and Grimm (2011), which are formulated within variants of Dowty's (1991) prototype-theory, the degree of transitivity is defined as the semantic distance between the co-arguments of a verb.

While the Affectedness Hierarchy only provides a 'unipolar' view of transitivity from the point of view of the object, it is nevertheless clear that a lower degree of affectedness correlates with decreased semantic distance between the co-arguments. Beavers (2010) also casts his approach within an entailment-based view of lexical meaning and he points out that his degrees of affectedness can be conceived of as corresponding to Dowty's Proto-Patient entailments; the relation between linking patterns of *-ung* nominalisations and the Affectedness Hierarchy may thus also be interpreted as the result of in-/decreasing transitivity caused by in-/decreasing semantic distance of the co-arguments. A lower level of affectedness thus corresponds to a less prototypical Proto-Patient which is semantically closer to the Proto-Agent and thus the verb is semantically less transitive. Since decreased transitivity may correlate with optional realisation of the verb's direct object (see the discussion of Levin, 1999, in section 1.5.2.1 above), the argument linking options with *-ung* nominalisations can be viewed as the consequence of different degrees of semantic transitivity: the odds that GenO vacates the postnominal genitive position in favour of GenS increase with reduced semantic distance between the two arguments of a syntactically transitive verb.

Recasting Ehrich and Rapp's approach to argument linking in eventive *-ung* nominalisations within Beavers' scalar model of affectedness potentially allows to capture more graded judgements and opens up some interesting alternative perspectives on the phenomenon: the semantic source of the observed linking patterns can be reconceived as semantic prominence induced by varying degrees of affectedness of GenO or as reduced/increased transitivity caused by in-/decreased semantic distance between the co-arguments. These two aspects are closely related and complementary and – in my view – they represent two sides of the same coin.



## 2.5 Excursus: on the syntactic status of postnominal genitives – arguments or adjuncts?

In this section, I provide a brief discussion regarding the status of postnominal genitives as arguments or adjuncts. Note that this section is not crucial for the understanding of the remainder of the thesis and may be skipped safely by linguistically less inclined readers.

While Ehrich and Rapp (2000), Bierwisch (1989) and others attribute argument status to postnominal GenO and GenS of eventive *-ung* nominalisations, others treat both or one of these as adjuncts (e.g. Bücking, 2010; Hartmann & Zimmermann, 2003; Solstad, 2010). The assessment of the syntactic status of GenO/GenS involves a number of different factors of which I will briefly address two. The first is the optional nature of genitives with German deverbal eventive (*-ung*) nominalisations; in her influential work, Grimshaw (1990) argues that ‘true’ complex event nominalisations take over the base verb’s argument structure, rendering the realisation of its argument (largely) obligatory (see section 2.1). Thus, German eventive nominalisations may actually be categorised as simple event nouns and GenO/GenS accompanying them as argument-adjuncts in the sense of Grimshaw, which are related to the verb only conceptually, but not grammatically.

With respect to German eventive nominalisations, Bierwisch (1989), Ehrich and Rapp (2000) and Ehrich (2003) do not assume that the facultative realisation of genitives goes hand in hand with adjunct status and Bierwisch (2009) rejects Grimshaw’s distinctions between CENs and SENs and thus also an analysis of GenO/GenS as argument-adjuncts. Ehrich (2003) uses plural nominalisations to argue for a syntactic argument status of GenS/GenO and against a conceptual relation between genitives and the underlying base verbs. Her argument goes as follows: while the nominal linking rules as stated in Ehrich and Rapp (2000) render linking of GenS illicit with *singular* nominalisations which embed a *BECOME* operator, the GenS interpretation becomes available along with the GenO reading when the same nominalisations are *pluralised* as in the examples in (87) (from Ehrich, 2003, p. 34). While this pattern actually seems to be in contradiction to the linking rules, Ehrich shows that eventive plural nominalisations actually behave like processes semantically, i.e. their LSS lacks *BECOME*; hence, the linking rules allow realisation of GenS and the behaviour of plural nominalisations actually strengthens Ehrich and Rapp’s nominal linking rules.

- (87) a. die Vergiftung des Apotheker- $s_{\text{GenO/*GenS}}$  / die Vergiftungen des  
the poisoning the.GEN pharmacist-GEN / the poisonings the.GEN  
Apotheker- $s_{\text{GenO/GenS}}$   
pharmacist-GEN  
‘The poisoning of the pharmacist / the poisonings of the pharmacist’  
b. die Zerstörung Rom- $s_{\text{GenO/*GenS}}$  / die Zerstörungen Rom- $s_{\text{GenO/GenS}}$   
the destruction Rome-GEN / the destructions Rome-GEN  
‘The destruction of Rome / the destructions of Rome’

On the other hand, the linking rules do not explain the genitive linking options found with other types of deverbal nominalisations, such as zero conversions or nominalised infinitives.

The latter, for example, do *not* allow linking of GenS even when the base verb denotes states or processes and thus lack *BECOME* as illustrated in (88) (from Ehrich, 2003, p. 32).

- (88) a. das Beobachten des Planet-en<sub>GenO</sub> / \*des Astronom-en<sub>GenS</sub>  
           the observing the.GEN planet-GEN / the.GEN astronomer-GEN  
           'the observing of the planet / of the astronomer'
- b. das Messen des Strom-s<sub>GenO</sub> / \*des Ingenieur-s<sub>GenS</sub>  
           the measuring the.GEN current-GEN / the.GEN engineer-GEN  
           'the measuring of the current / of the engineer'

Thus, Ehrich (2003) concludes that the nominal linking rules formulated by Ehrich and Rapp are part of the *grammar* of *-ung*-nominalisations, but do not cover other types of nominalisations. If GenO/GenS were argument-adjuncts and their linking guided by conceptual rather than grammatical processes, the derivation type should not matter – the fact that it does underscores the grammatical source of the linking processes with *-ung*-nominalisations and the argument status of GenO/GenS.

The second issue I will address pertains to an asymmetry in the interpretive (in)variance of GenO and GenS. The issues involved here are both subtle and complex and a thorough discussion is not my aim – an in-depth discussion can be found in Bücking (2011, section 4.4). At the heart of the issue lies the fact that genitives accompanying non-eventive nouns can in principle have a number of different interpretations: in *das Buch Nellas* (*Nella's book*), *Nella* may be understood to be the author, the buyer, a fervent reader, the possessor of the book, or to stand in a number of other possible relations to the book. Such variable interpretive possibilities are indicative of rather free relations not guided by argument/thematic structure. Bücking (2010) observes an interpretive asymmetry between postnominal GenS and GenO focussing on nominalised infinitives as in (88) and claims that GenS shows more interpretive freedom than GenO, with the latter always being interpreted as the thematic role underlying the base verb's direct object. He concludes that this asymmetry speaks for an argument status of GenOs, while GenSs are treated as adjuncts related to the base verb solely by a 'malleable' conceptual relation which – among others – also allows for Agent-*like* interpretations such as Possessor, which I will refer to as 'free genitive (interpretations)' in the following.

While I take linking patterns observed with nominalised infinitives to be of limited evidence for aspects of argument linking related to *-ung*-nominalisations in the light of the conclusions of Ehrich (2003) discussed above, the issue can be illustrated with *-ung*-nominalisations; some *-ung*-nominalisations like *Beschreibung/description* exhibit a semantic ambiguity and can be interpreted as events or the respective outcome, which denotes a concrete, physical object, i.e. a simple (non-relational) noun. In (89a) (example a. is from Bücking, 2011, p. 163, the others are mine), the physical object interpretation is forced by the context and *Pauls* can accordingly be interpreted in a number of ways, including free genitive readings such as author or possessor; *des Buches/of the book*, on the other hand, still receives the its interpretation as determined by the verb's LSS (corresponding to Theme in traditional thematic role

approaches). (89b) shows that an LSS-bound GenS interpretation – which is forced here by the adjective *halbstündig/half-hour-long* – is not compatible with a physical object reading of *Beschreibung/description*; vice versa, (89c) shows that an free genitive interpretation of *Pauls* (forced by *zerknittert/crumpled*) is not reconcilable with the eventive interpretation of *Beschreibung/description* (which is forced, in turn, by the verb *unterbrechen/interrupt*).

- (89) a. Paul-s<sub>Author/Possessor/...</sub> Beschreibung des Buch-es<sub>Theme</sub> liegt auf dem  
Paul-GEN description the.GEN book-GEN lies on the  
Tisch.  
table  
'Paul's description of the book lies on the table.'
- b. \* Paul-s<sub>GenS</sub> halbstündige Beschreibung des Buch-es liegt auf dem  
Paul-GEN half-hour-long description the.GEN book-GEN lies on the  
Tisch.  
table  
'Paul's half-hour-long description of the book lies on the table.'
- c. \* Der Lehrer unterbrach Paul-s<sub>Author/Possessor</sub> zerknitterte  
the.NOM teacher.NOM interrupted Paul-GEN crumpled  
Beschreibung des Buch-es.  
description the.GEN book-GEN  
'The teacher interrupted Paul's crumpled description of the book.'

In my opinion, while the above examples illustrate that free genitive interpretations are of course available with *-ung*-nominalisations in certain contexts (for example when the nominalisation denotes a concrete object), they also show that the eventive and other interpretations of ambiguous nominalisations are mutually exclusive – consequently, provided that the nominalisation itself is disambiguated by the context, the interpretation of the genitive as either LSS-bound GenS on the one hand or a free genitive on the other hand should be disambiguated too. As Ehrich and Rapp (2000) point out, while the genuine GenS and free genitive (among which is the *genitivus auctoris*, which refers to the creator of an object) interpretations are conceptually related, they are not the same. They can be syntactically differentiated by the fact that a 'true' GenS can be replaced by a PP-adjunct headed by *durch* (91a), but a free genitive cannot (91b):

- (90) a. Der Lehrer unterbrach die Beschreibung des  
the.NOM teacher.NOM interrupted the.ACC description.ACC the.GEN  
Buch-es durch Paul<sub>Agent</sub>.  
book-GEN by Paul  
'The teacher interrupted the book's description by Paul.'
- b. \* Die Beschreibung des Buch-es durch Paul<sub>Author/Possessor/...</sub>  
the.NOM description.NOM the.GEN book-GEN by Paul  
liegt auf dem Tisch.  
lies on the table  
'The book's description by Paul lies on the table.'

Other tests used by Ehrich and Rapp to identify eventive readings of deverbal nominalisations include compatibility with verbs and adjectives which force the process or event interpretation. The fact that *Behandlung/treatment* (which only allows for a process reading) allows for modification by *langwierig/prolonged* and can occur as the object of *unterbrechen/interrupt* with either GenO or GenS realised as postnominal genitive in (91) underscores the compatibility of an eventive reading of the nominalisation with linking of GenS and thus the genuine argument character of GenS in this context.

- (91) a. Die Angehörigen unterbrachen die langwierige Behandlung  
the.NOM relatives.NOM interrupted the.ACC prolonged treatment.ACC  
des Patient-en<sub>GenO</sub>.  
the.GEN patient-GEN  
'The relatives interrupted the prolonged treatment of the patient.'
- b. Die Angehörigen unterbrachen die langwierige Behandlung  
the.NOM relatives.NOM interrupted the.ACC prolonged treatment.ACC  
des Arzt-es<sub>GenS</sub>.  
the.GEN doctor-GEN  
'The relatives interrupted the prolonged treatment by the doctor.'

Although a vague context can of course render a nominalisation as well as an accompanying genitive ambiguous, I believe the above examples show that a genuine argument interpretation of GenS is available. In the remainder of this thesis, I will thus follow Ehrich and Rapp (2000) and will assume that genuine GenS interpretations are available with eventive *-ung*-nominalisations and that these are anchored to the verb's semantic structure via argument structure, just as their GenO counterparts. GenS may of course receive free genitive interpretations where the necessary non-eventive interpretation of the nominalisation is licensed by the context.<sup>7</sup>

## 2.6 Eventive nominalisations as a backdoor to affectedness-related linking effects at the syntax/semantics interface

Having reviewed relevant aspects of nominal linking and the semantic determinants of argument linking with eventive *-ung*-nominalisations in detail, I will justify the use of *nominals* for the exploration of the semantic basis of *verbal* argument linking; after all, the route via nominalisations certainly constitutes a major detour which calls for explanation.

First, as reviewed above, eventive nominalisations inherit the base verb's lexical-semantic structure as well as (parts of) its argument structure. Since these two together form the

<sup>7</sup>Bücking (2010, p. 52) and Bücking (2011, p. 161) point out that the free genitive interpretation is actually often marked and requires special contexts, at least with nominalised infinitives.

relevant aspect of the syntax/semantics interface, nominalisations *can* in principle be used to investigate the questions which this thesis aims at addressing. But why *should* they?

Whereas the parallels between verbs and their eventive nominalisations provide the basis for their application in the experimental investigation of verbal argument linking, in my view it is the *differences* between verbal and nominal linking processes which motivate and justify it. In German *-ung*-nominalisations, the specific nominal argument linking processes and the particular restrictions imposed by the structure and morphosyntax of German NPs together result in a constellation which provides a unique testing ground for the questions of interest: first, following Ehrich and Rapp (2000) the postnominal position for genitives in German NPs is the only 'true' argument position available for linking of GenO and GenS, which results in a competition of the two arguments of a syntactically transitive verb for this slot. Second, also following Ehrich and Rapp the factors determining which of the two arguments may win this competition with *-ung*-nominalisations are purely semantic in nature, since the relevant information born by the arguments goes back to the verb's semantic structure.

I argued above that Ehrich and Rapp's analysis of the linking rules relevant for *-ung*-nominalisations may be reinterpreted within the scalar approach to degrees of affectedness of Beavers (2010, 2011), which potentially allows to capture more graded acceptability judgements about the linking possibilities of GenO/GenS. Thus, *judging whether a specific eventive -ung-nominal licenses linking of GenO only or of GenS as well requires: a) a check of the arguments' predicate-induced semantic properties, especially the level of affectedness entailed of GenO and b) processing this information further to regulate the linking competition between the co-arguments*. Both of these are key aspects of the central research questions formulated in the Introduction.

In addition, viewing Ehrich and Rapp's nominal linking rules as reflexes of degrees of affectedness opens up a number of complementary conceptual perspectives: it allows to interpret the patterns as determined by semantic prominence of the direct object participant induced by verbal semantics and provides links to prototype-based linking approaches and a view of semantic transitivity which makes recourse to a notion of semantic distance between co-arguments of a verb (see section 2.4). Checking the relevant predicate-induced semantic property of the arguments can thus be understood as accessing semantic prominence information about the participant linked as the direct object and/or about the semantic distance between the respective co-arguments, both as defined by degree of affectedness.

Crucially, *with eventive -ung-nominalisations checking the relevant predicate-induced semantic argument property is the ONLY cue available to resolve the competition between GenO and GenS, since all morphosyntactic cues potentially available in the verbal domain are neutralised*: the morphosyntactic processes between the head noun of an NP and arguments related to it are limited to the assignment of genitive case in German, while other processes such as person/number agreement between them are lacking. Further, GenS and GenO are *both* marked by genitive case. Compared to the verbal domain, morphosyntactic processes between the predicative head and its arguments are thus drastically reduced, since in German verbs show morphological agreement with their subjects in person and number and subjects and direct objects are differentially marked with nominative and accusative case, respectively.

Consequently, ambiguities about the status of arguments of verbs can be resolved by morphological agreement; in example (92a), the masculine singular NPs *der Mann/the man* and *den Jungen/the boy* are unambiguously marked with nominative and accusative case, respectively, and the verb agrees with the sentence-initial subject NP *der Mann* in person and number. Due to morphological syncretism and flexibility in the relative order of subject and direct object in German, the feminine singular NPs *die Mutter/the mother* and *die Tochter/the daughter* in (92b) render this sentence ambiguous: while the preferred interpretation is subject/object, the sentence final NP *die Tochter* can in principle be the subject, resulting in non-canonical object/subject word order. In (92c) the less preferred object/subject reading is forced by the morphological number agreement between the verb and the sentence-final plural NP *die Töchter/the daughters*; while both argument NPs are in principle still locally ambiguous with respect to their syntactic status, the linking is disambiguated by the agreement between the verb and the subject.

- (92) a. Der Mann sah den Jungen.  
           the.MASC.SG.NOM man saw.PST.3SG the.MASC.SG.ACC boy  
           ‘The man saw the boy.’
- b. Die Mutter sah die Tochter.  
           the.FEM.SG.NOM/ACC mother saw.PST.3SG the.FEM.SG.NOM/ACC daughter  
           ‘The mother saw the daughter./The daughter saw the mother.’
- c. Die Mutter sah-en die Töchter.  
           the.FEM.SG.NOM/ACC mother saw.PST-3PL the.FEM.PL.NOM/ACC daughters  
           ‘The daughters saw the mother.’

Thus, ambiguities about the status of co-arguments in the verbal sentence domain can be resolved via morphosyntactic processes between the verbal head and its arguments and it is not clear to what degree the parsing may in such cases be influenced lexical-semantic properties of the respective verb; in my view, for example, replacing the semantically minimally transitive verb *saw* with the maximally transitive verb *killed* in (92b) or (92c) does not make any noticeable difference with respect to the intuitions about interpretive preferences or ease of reanalysis. While it is possible that parsing processes in these cases are subtly influenced by the verb’s entailments about degrees of affectedness, for example, the influence of affectedness for nominal argument linking is immediately appreciable. Thus, by using eventive nominalisations to probe for linking-related processing correlates at the syntax/semantics interface (via affectedness), the morphosyntactic route can be completely by-passed for resolving any linking conflicts, minimising possible overlapping effects due to processes operating on verbal semantic structure and its reflexes on argument NPs on the one hand and on morphosyntax on the other. The rationale for taking the detour via nominalisations is hence that this alternate route may actually provide a direct shortcut to the processes mediating between verbal semantics and argument linking.

Whereas the use of nominalisations allows avoidance of morphosyntactic processes for the

resolution of linking conflicts, it also comes at a cost: since an isolated phrase like *die Bewunderung des Jungen/the admiration of the boy* is in principle ambiguous between a GenO and GenS interpretation, disambiguation of the argument status of *des Jungen* requires prior contextual specification. In (93a), the status of the two verbal arguments is made clear, so *des Jungen* in (93b) can only be interpreted as GenS, whereas *des Großvaters/of the grandfather* in (93c) is restricted to be understood as GenO. This disambiguation involves accessing the discourse representation established by (93a) and may thus result in the tapping of resources related to the processing of discourse context (see section 3.2.6 below for a brief overview of relevant aspects).

(93) a. *Context:*

Der Junge bewunderte den Großvater.  
the.NOM boy.NOM admired the.ACC grandfather.ACC

'The boy admired the grandfather.'

b. die Bewunderung des Junge-n\*<sub>GenO/GenS</sub>  
the admiration the.GEN boy-GEN  
'the admiration of the boy'

c. die Bewunderung des Großvater-s\*<sub>GenO/\*GenS</sub>  
the admiration the.GEN grandfather-GEN  
'the admiration of the grandfather'

Another inherent risk of the present approach lies in the danger of actually investigating processes which are restricted to a niche context, i.e. the specific environment of eventive (*-ung*) nominalisations made use of here. In fact, this thesis does not provide experimental support that any processing correlates identified generalise to online linking processes within the verbal, sentential domain. Thus, the current investigation partially has an exploratory character and any processing correlates potentially identified will have to be validated by independent experiments within the verbal domain, a point addressed further in the concluding discussion in chapter 5. In fact, one of the central aims is to provide a starting point for formulating hypotheses about possible processing effects at the syntax/semantics interface with verbal argument linking.

## 2.7 Acceptability judgement experiment

Above, I have reviewed the interaction of the semantic property of degree of affectedness and argument linking patterns associated with postnominal genitive arguments in German eventive *-ung*-nominalisations. In the previous section, I presented arguments that the use of deverbal, eventive nominalisations, rather than verbs, may actually constitute a shortcut for investigating the role of affectedness in online argument linking processes. In addition, I proposed that the Affectedness Hierarchy developed by Beavers (2010, 2011) may allow to account for more graded intuitions about the acceptability of linking GenS arguments, in

contrast to the approach to argument linking with *-ung*-nominalisations presented by Ehrich and Rapp (2000), which links acceptability of GenS linking to the absence or presence of a *BECOME* operator in the base verb's structural semantic representation.

In the following sections, I will present an acceptability judgement experiment which investigates if the acceptability of postnominal genitive linking in German *-ung*-nominalisations indeed follows a graded pattern which is compatible with the notion that affectedness is a matter of degree, as stated by Beavers' Affectedness Hierarchy. This initial experiment had the following aims: first, it was intended to validate this claim and obtain empirical evidence for the influence of affectedness for GenS as well as GenO linking. It thus assesses my mapping of Beavers' Affectedness Hierarchy onto the approach to argument linking in eventive *-ung* nominalisations developed by Ehrich and Rapp (2000), as presented in section 2.4. If this interpretation receives empirical support, the linking patterns may be thought of as a consequence of the semantic prominence of GenO – as determined by varying degrees of affectedness – or, on a complementary view, as varying semantic transitivity determined by the semantic distance between the GenO and GenS co-arguments. The second aim was to validate the experimental paradigm, which will provide the basis for the subsequent online processing experiments.

Note that in the experiments to follow, for practical reasons I implemented the concept of degrees of affectedness as formulated by Beavers (2010, 2011) as a continuous variable, rather than as a variable with discrete levels. I will discuss the rationale for thus re-operationalising degrees of affectedness in the following section. Section 2.7.2 presents a pretest in which ratings about the degree of affectedness implied by a set of verbs were obtained from participants for the construction of experimental materials. Section 2.7.3 then presents the acceptability judgement study, with the specific hypotheses about the acceptability patterns provided in section 2.7.3.3.

### 2.7.1 Operationalising degree of affectedness as a continuous variable

In section 2.4 above, I suggested to reinterpret Ehrich and Rapp's argument linking approach in eventive *-ung* nominalisations in terms of scalar model of affectedness developed by Beavers (2010, 2011); on this view, the nominal linking patterns can be conceived of as the consequence of the semantic prominence of GenO as defined by the Affectedness Hierarchy. This hierarchy comprises the four well defined categorical levels given in (49) (*unspecified for change*; *potential for change*; *non-quantised change*; *quantised change*) and the examples given in (86) suggest that these four degrees may well capture the acceptability of relevant linking patterns in eventive *-ung* nominalisations.

One difficulty with applying this four-level Affectedness Hierarchy for empirical experiments in the current context lies in the construction of stimuli given a certain degree of uncertainty about the classification of verbs which cannot be clearly assigned to a given level on the hierarchy. As an example, take the verb *touch*, which is classified as belonging to the lowest level of the Affectedness Hierarchy, i.e. as unspecified for change, by Beavers (2010, 2011).



Verbs which are unspecified for change can be distinguished by force recipient verbs one level higher (which show potential for change) by the *What happened to X is Y* entailment text (see section 1.5.1 in the previous chapter), which should only be passed by force recipient verbs (or verbs higher up the hierarchy). Beavers (2010, p. 836) gives the following example (the judgement is his):

(94) ?? What happened to the wall is that John touched/grazed/saw it.

In his footnote 10, Beavers (2010, p. 836) mentions that “[s]uch predicates are acceptable in *What happened to X is Y* in very particularized contexts, for example, a King Midas context where a touch turns something to gold” – however, since this is licensed by the context and not by verbal entailments, Beavers argues that such factors are irrelevant to the classification, albeit he acknowledges that they will “muddle the judgments somewhat”. Example (95), however, illustrates that it takes much less than a King Midas context to make verbs such as *touch* or *graze* compatible with the *What happened to X is Y* test: here, it is the shaky nature of the object NP fillers which makes the example acceptable.

(95) What happened to the house of cards/tower of toy blocks is that Luk touched/grazed it.

The influence of the lexical filler of the direct object NP can also be seen with some verbs belonging to higher levels of the Affectedness Hierarchy. As an illustration, when the verb *disarm* takes an object like *the bomb*, it denotes quantised change, since the potentially relevant states of the bomb only comprise two levels – armed or not armed. However, when it takes an NP like *the situation* as object, for example, it is not clear whether the change implied is quantised or non-quantised; after all, one can disarm a situation completely or only somewhat.

Since the pool of potential items is already rather limited given the criteria listed in the following section, such difficulties may reduce the number of candidates further and/or may result in problems finding sufficient items for some of the four affectedness levels. Thus, I will use a different operationalisation of degree of affectedness which avoids these issues, but is nevertheless based upon Beavers’ Affectedness Hierarchy and its main insights and should thus be able to capture potential effects in a similar way: rather than implementing the four discrete levels of the hierarchy as a four-level variable, degree of affectedness will be operationalised as a continuous numerical covariate, the ‘Affectedness Index’ (henceforth AI). Thus, the degree of affectedness implied by a given base verb (and thus of a corresponding nominalisation) will be represented by a numerical value, rather than by membership in a discrete category. The AI for each base verb/nominalisation was acquired in a pretest in which naive German speakers rated the degree of affectedness implied for the direct object of a verb; since these ratings are based solely upon the degree of change the object referents *necessarily* undergo during the event described (see section 2.7.2.2 below), the resulting AI should be able to roughly capture hierarchical relations between verbs/nominalisations based upon the degree of affectedness and the associated semantic prominence in a way

similar to Beavers' Affectedness Hierarchy (see also Hindy, Altmann, Kalenik, & Thompson-Schill, 2012, for a similar procedure for acquiring continuous change-of-state indices in an fMRI study of object state-change).

At the same time, however, it is apparent that by undertaking this conversion one runs the risk of reintroducing a certain degree of 'fuzziness' into the determination of the degree of affectedness entailed by a predicate, which Beavers explicitly sought to eliminate by separating the scale encoding the degree of affectedness from real-world related aspects; after all, speakers' intuitions about the degree of change implied by a verb may vary due to a number of sources, including the concrete sentential context used in the AI-pretest study. In this respect, the AI used in the subsequent experiments is on a par with a number of other variables which are based on ratings or other responses given by naive speakers, including cloze probabilities, imageability or familiarity ratings. All of these have nevertheless been successfully employed in a vast number of psycholinguistic studies.

The analysis of the experimental data will be conducted using linear mixed effects models (or 'linear mixed models' – henceforth LMMs), a regression-based analysis tool which allows to include continuous numerical variables in an analysis without the need to resort to dichotomising them. LMMs provide a number of additional advantages to more traditional ANOVA-based analyses, such as the ability to handle binomial data and subject- as well as item-related random effects in one and the same model; for discussion, see Baayen et al. (2008), Jaeger (2008) as well as section 6.2 in the second part of this thesis. On my view, the inclusion of 'AI' as a continuous covariate into the LMM constitutes a conceptual simplification of the resulting model when compared to a model with a categorical (ordinal) four-level variable: rather than interpreting the effects of the separate levels of the categorical variable along with its interaction with the two-level variable 'Linking' (with levels 'GenO' and 'GenS'), it suffices to assess the slope of the continuous 'AI' predictor and its interaction with 'Linking'. Note that such a model allows to adequately assess the predictions formulated in section 2.7.3.3 below.

## **2.7.2 Pretest**

The pretest presented in the following aimed at obtaining AI values for a set of verbs for the construction of experimental materials. To this end, ratings about the degree of affectedness implied by syntactically transitive verbs for their direct object were obtained from naive participants.

### **2.7.2.1 Materials**

From an initial list of German nouns ending in *-ung* assembled from the Celex data base (Baayen, Piepenbrock, & van Rijn, 1993) and other sources, a set of *-ung* nominalisations was selected according to the following criteria:

- Each nominalisation's base verb is syntactically transitive, i.e. subject and object must be realised with the unmarked NOM/ACC pattern. In addition, the base verb allows

the formation of a verbal passive with the auxiliary *werden*. These criteria result in the inclusion of syntactically strongly transitive verbs (see section 1.5.2.1 in the previous chapter).

- The nominalisation has an eventive reading corresponding to its base verb’s most prominent aspectual interpretation (e.g. state, process, achievement or accomplishment). Nominalisations which have lexicalised meanings and cannot be interpreted eventively (e.g. *Heizung/heating*, which only allows the concrete object reading corresponding to *radiator*) were excluded. So were nominalisations which have prominent result object interpretations, since these may allow the linking of genitives with free genitive interpretations such as Possessor (see section 2.5).
- The base verb allows the subject as well as the direct object slot to be filled by NPs headed by nouns denoting humans or human-like entities (e.g. animals or collective nouns). This criterion was included to avoid potential confounds of effects related to the linking of GenO/GenS on the one hand and differences in the animacy of lexical fillers for GenO/GenS on the other hand in the processing experiments presented in the following chapters.

Thus, the selected nominals are all syntactically strongly transitive, but vary along the conceptual and semantic dimensions; the resulting set of 85 nominalisations comprised items from a number of different ‘real world’ conceptual-semantic fields, i.e. the types of the scales (in the sense of Beavers, 2011) associated with the respective base verb varied and included location scales (e.g. *verschleppen/verlegen/einliefern* (*carry off, abduct/transfer, relocate/hospitalise*)), scales based upon different kinds of physical (e.g. *behandeln/untersuchen/durchsuchen/heilen/töten* (*treat/examine/frisk/cure/kill*)) and mental processes (e.g. *bewundern/belehren/einschüchtern/aufmuntern* (*admire/instruct, mirandise/intimidate/cheer up*)), status scales (e.g. *befördern/rehabilitieren/verurteilen* (*promote/rehabilitate/convict*)) as well as mixed scales incorporating different aspects of these and/or additional types. Also, the base verbs of the resulting set included verbs of all degrees of affectedness posited by Beavers, including verbs unspecified for change such as *bewundern/bewachen/überwachen* (*admire/guard/observe*), verbs which are associated with potential for change like *behandeln/durchsuchen/untersuchen* (*treat/frisk/examine*), non-quantised change (e.g. *aufmuntern/befördern* – *cheer up/promote*) and verbs of quantised change like *töten/entlassen/einbürgern* (*kill/dismiss/naturalise*). The nominalisations within the selected set are derived from base verbs with morphologically heterogeneous structure, comprising words without prefixes such as *töten/heilen/fördern* (*kill/cure/sponsor*) and with a number of different prefixes like *an-hören/er-pressen/ver-bannen* (*hear (a party)/blackmail/ban*).

To acquire ratings about the AI of the selected nominalisations, short transitive subject-verb-object (SVO) sentences were constructed using the nominalisations’ base verbs as verbs and – wherever possible – singular nouns of masculine gender denoting human entities as heads of the respective subject and object NPs. NPs headed by masculine singular nouns bear unambiguous nominative/accusative case marking, maximally disambiguating the arguments’ grammatical status. In few cases, neuter nouns were used; in these cases, the other NP

was unambiguously marked with nominative or accusative, however, still allowing for easy and unambiguous identification of subject and object. Further, in exceptional cases collective nouns such as *Senat/Vorstand* (*senate/board of directors*) or other non-human, animate nouns such as *Tiger/Löwe* (*tiger/lion*) were used as lexical fillers for the subject or object NP. The verbs were inflected for past tense (*Präteritum*). (96) lists a number of example sentences.

- (96) a. Der Sohn bewunderte den Star.  
the.NOM boy.NOM admired the.ACC star.ACC  
'The boy admired the star.'
- b. Der Agent überwachte den Politiker.  
the.NOM agent.NOM watched the.ACC politician.ACC  
'The agent watched the politician.'
- c. Der Arzt behandelte den Patienten.  
the.NOM doctor.NOM treated the.ACC patient.ACC  
'The doctor treated the patient.'
- d. Der Chef versetzte den Mitarbeiter.  
the.NOM boss.NOM relocated the.ACC employee.ACC  
'The boss relocated the employee.'
- e. Der Wilderer tötete den Löwen.  
the.NOM poacher.NOM killed the.ACC lion.ACC  
'The poacher killed the lion.'

To reduce possible effects of specific lexical fillers of subject/object NPs on the AI ratings, two such sentences were constructed for each base verb in the set, resulting in a total of 170 stimulus sentences. These were distributed over two lists of 85 sentences each, with each list featuring each verb once. The order of presentation of stimuli within a list was randomised for each participant.

### 2.7.2.2 Procedure

The AI-ratings were collected using WebExp, a software toolbox developed for conducting psychological experiments via the internet (Keller, Corley, Corley, Konieczny, & Todorascu, 1998), allowing the participants to access and rate the stimulus sentences remotely via a web-browser. On the first page, the participants were informed about the general institutional background of the experiment and use of the data collected. They were informed that they would be asked to indicate general personal information such as age and sex, but that they would not be asked for their names and that the experiment would be analysed anonymously. In addition, participants were told that only native speakers of German were eligible for participation in the study.

On the following page, participants received detailed instructions about the experimental task; they were informed that during the experiment they would be shown one short sentence

at a time for which to provide the ratings.<sup>8</sup> To ensure easy identification of the participant realised as the direct object in the short SVO sentences, the object NP was highlighted by underlining in each stimulus sentence. Participants were instructed to provide ratings about the degree to which the highlighted event participant underwent a change in the course of the event described by the respective sentence. As an orientation, they were told to ask the following question for each rating: if the event took place as described, did the highlighted participant **NECESSARILY** undergo a change when compared to the time before the event? Participants were instructed to provide their ratings on an integer scale from 1–7, with 1 indicating that the relevant participant definitely did not undergo any kind of change (or that it is left completely open whether it did) and 7 indicating that it definitely underwent a specific change.

Participants were informed that they should consider a number of different types of change, such as change in physical and mental state, legal, occupational or social status, change in location and others. The instructions contained three commented example sentences as well as examples for each type of change. Participants were instructed to base the ratings on their intuitions, follow their gut feeling, provide the ratings swiftly and to use the whole range of the rating scale.

### 2.7.2.3 Participants

Complete ratings were collected from 26 subjects; the data of one of these subjects had to be discarded due to uncooperative behaviour, leaving 25 participants with a mean age of 21.6 (range 19–30, 1 male, 24 female) which entered the final analysis. Note that a number of additional participants took part in the rating study, but due to unknown technical problems, their ratings were not saved upon completion of the questionnaire. The data of these participants could thus not be considered. Subjects were assigned to one of the two experimental lists on a random basis; 13 of the participants included for analysis were attributed list 1 and the remaining 12 to list 2. For participation in the study, participants could opt for course credits or for taking part in the drawing of a voucher for an online store.

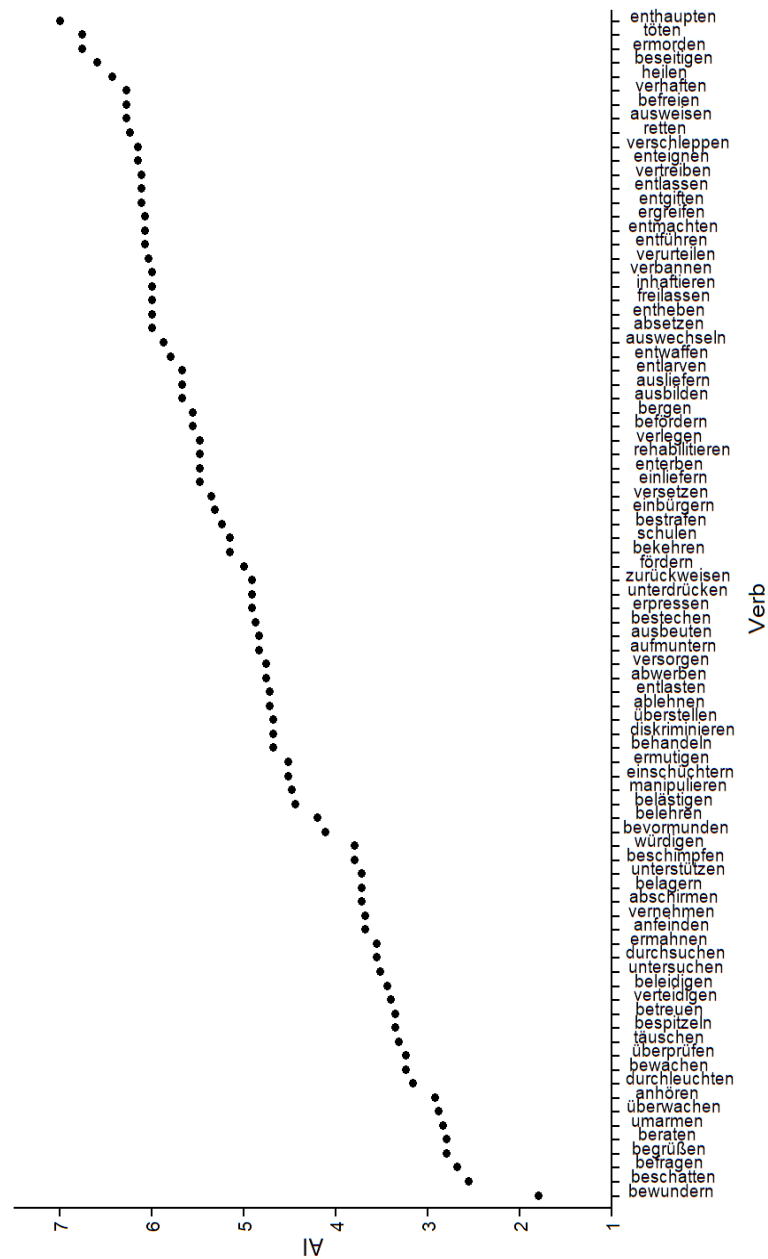
### 2.7.2.4 Results

The AI of a given verb was calculated as the mean of the object-related change ratings across all subjects. Figure 2.1 shows the AI for each verb included in ascending order: the lowest AI of 1.8 is attributed to the verb *bewundern/admire* and the highest AI of 7 to the verb *enthaupten/decapitate* and the verbs in between cover the whole range of the scale in a (nearly) continuous way with a mean AI of 4.82 (median: 4.92, SD: 1.23).

The five verbs with the lowest AI are *bewundern/beschatten/befragen/begrüßen/beraten* (*admire/tail/interview, interrogate/greet/advise*), all of which indeed do not necessarily entail any kind of change in the participant realised as object and can be associated with

---

<sup>8</sup>In addition to the rating about AI, participants provided two further ratings for each sentence. Since these do not play a role in the remainder of this study, I will not go into any details about related instructions.



**Figure 2.1:** AI rating for each verb (sorted by AI in ascending order; figure has been rotated by 90 degrees).

the lowest degrees of Beavers' Affectedness Hierarchy. The five verbs with the highest

AI are *enthaupten/töten/ermorden/beseitigen/heilen*<sup>9</sup> (*decapitate/kill/murder/remove/cure*), which all entail a concrete and definite change in the object participant and belong to the highest level of the Affectedness Hierarchy, quantised change. While verbs in the mid-range are more mixed, the five verbs around the median AI of 4.92 – *ausbeuten/besteichen/erpressen/unterdrücken/zurückweisen* (*cheer up/exploit/bribe/blackmail/oppress, subdue*) – are largely consistent with classification as ‘potential for change’ or ‘non-quantised change’ predicates, i.e. the two levels covering the middle of the Affectedness Hierarchy. Thus, the continuous AI seems to reasonably capture the degree of affectedness entailed for a verb’s direct object and thus also its semantic prominence as conceptualised by Beavers (2010, 2011).

## 2.7.3 Acceptability judgement experiment

### 2.7.3.1 Materials

The *-ung*-nominalisations of the 85 base verbs (see also table A.1 in Appendix A for details on the nominalisations used) from the pretest were used to create materials for eliciting acceptability judgements about the possibility of linking GenS and GenO with each of these. To do so, sentence pairs were constructed for each nominalisation; these expanded the stimulus sentences applied in the pretest into short ‘stories’, as in examples (97) and (98): the first sentence (henceforth ‘context sentence’ – see (97a) and (98a)) always set up the discourse context by describing an event or situation involving two participants. These corresponded to the respective base verb’s subject and object NPs used in the AI-rating pretest or were selected according to the same criteria stated above. In this context sentence, the base verb of the respective nominalisation was not used directly; rather synonyms of the respective verb or paraphrases were used. The second sentence (henceforth ‘continuation sentence’ – see (97b) and (98b)) represented a continuation of the story introduced by the context sentence; it featured the respective *-ung*-nominalisation as head of the sentence-initial subject NP. The nominalisation was followed by a post-nominal genitive representing the subject of the context sentence (i.e. GenS) or its object (i.e. GenO). The whole subject NP containing the nominalisation and the postnominal genitive was followed by the main verb inflected for past tense and further sentence material completing the sentence.

(97) Example story for *Bewunderung* (*admiration*):

a. *Context sentence with ‘role allocation’:*

Der Sohn himmelte den Star Tag und Nacht an.  
the.NOM son.NOM adored the.ACC star.ACC day and night PART

<sup>9</sup>Note that *heilen* (*cure, heal*) can occur as a syntactically intransitive ‘unaccusative’ verb – however, in this usage the subject slot must be filled by an entity which does not denote a human or animal, such as a body part as in *der Arm heilte/the arm healed*, whereas human subjects are unacceptable as in *\*der Patient heilte/the patient healed*. Since subject and object NPs always referred to humans in this and all following experiments, it is clear that *heilen* here is used in its syntactically transitive form. Note that Blume (2000, p. 181) also lists *heilen* as a typical case of a semantically transitive verb.

'The son adored the star day and night.'

b. *Continuation sentence*:

Die Bewunderung des Star-s<sub>GenO</sub> / des Sohn-es<sub>GenS</sub> beunruhigte den  
the admiration the.GEN star-GEN / the.GEN boy-GEN worried the  
Vater nach einiger Zeit.  
father after some time

'The admiration of the star/of the boy worried the father after some time.'

(98) Example story for *Ermordung* (murder/assassination):

a. *Context sentence with 'role allocation'*:

Der Räuber erschoss den Wächter während seiner Flucht aus  
the.NOM robber.NOM shot the.ACC guard.ACC during his flight from  
der Bank.  
the bank

'The robber shot the guard during his flight from the bank.'

b. *Continuation sentence*:

Die Ermordung des Wächter-s<sub>GenO</sub> / des Räuber-s<sub>GenS</sub> schockierte  
the assassination the.GEN guard-GEN / the.GEN robber-GEN shocked  
die Angestellten in der Bank.  
the employees in the bank

'The assassination of the guard/of the robber shocked the bank's employees.'

For each of the 85 nominalisations, two different stories were constructed, resulting in a total of 170 stories. Each of these, in turn, was paired with either GenO or GenS as an argument to the respective *-ung*-nominalisation in its continuation sentence, yielding a pool of 340 sentence pair stimuli. These were distributed over four lists: each list contained 85 stories, one for each nominalisation; half of these featured GenS in the continuation sentence (condition 'GenS') and the other half GenO (condition 'GenO'), with either of these conditions covering the range of the AI representatively. Subjects were assigned to one of the four experimental lists on a random basis and the number of subjects per list was balanced. The order of presentation of stimuli within a list was again randomised for each participant. The complete set of stimuli is provided in Appendix B.

### 2.7.3.2 Procedure

The acceptability ratings were again collected via the internet using the WebExp software of Keller et al. (1998). On the first page, participants were informed about the general institutional background of the experiment and use of the data collected. They were informed that they would be asked to indicate general personal information such as age and sex, but that they would not be asked for their names and that the experiment would be analysed anonymously. In addition, participants were told that only native speakers of German were eligible for participation in the study.



During the experiment, each story was presented on a separate page, with the context sentence printed in black font and the continuation sentence in green font. Participants were told that they would read one pair of sentences on each page and that the first (black) sentence introduced a certain situation or event with two participants. They were further informed that the first sentence would define the 'role' each of the participants played in this event. The participants were then told that the second (green) sentence was intended to represent a continuation of the event/situation set up by the first one and that only *one* of the two participants introduced would appear in this continuation sentence. The subjects were instructed to judge whether the participant appearing in the continuation sentence still had the same role as in the context sentence – and thus could act as a coherent continuation of the event/situation introduced by the first sentence – or not. The judgement was provided by selecting 'Yes' or 'No' from a box titled 'Continuation'.

Since the GenO linking is assumed to be the preferred linking option throughout and occurs much more frequently in natural speech, the participants were asked to base their judgements on considering whether the second sentence could in principle act as a continuation, even if accepting this interpretation took some mental effort – if so, they should select 'Yes', if not 'No'. They were also informed that there were no 'correct' or 'incorrect' answers and they were instructed to follow their own linguistic gut feeling in providing the judgements and to provide their responses quickly. The instructions contained two commented example sentence pairs. After providing judgements about four example stories, the experimental stimuli were presented page by page.

### 2.7.3.3 Hypotheses

As outlined at the beginning of section 2.7.3, the primary aim of the acceptability judgement experiment was to test whether the acceptability of GenS linking in eventive *-ung*-nominalisations is indeed graded, as I conjectured in section 2.4. There, I proposed that a graded notion of affectedness as developed by Beavers (2010, 2011) may capture intuitions more adequately than the approach of Ehrich and Rapp (2000). If empirically supported, the linking mechanism investigated here may indeed be conceptualised as being semantically determined and a result of the semantic prominence of GenO, as indexed by its degree of affectedness. As I pointed out, a closely related interpretation involves the degree of semantic transitivity as determined by the semantic distance between the GenO and GenS co-arguments.

The examples in (86), based upon my own intuitions and repeated here, illustrated the crucial pattern mapped onto Beavers' Affectedness Hierarchy, with acceptability of GenS linking decreasing from a maximum to a minimum as affectedness increases, while GenO linking always seems equally acceptable.

(86) a. *Unspecified for change:*

die Bewunderung	/	Begrüßung	✓✓des	Patient-en <sub>GenO</sub>	/	✓✓des	Arzt-es <sub>GenS</sub>
the admiration	/	greeting	the.GEN patient-GEN	/	the.GEN	doctor-GEN	

'the admiration/greeting of the patient/the doctor'

b. *Potential for change:*

die Behandlung / Untersuchung ✓✓des Patient-en<sub>GenO</sub> / ✓des Arzt-es<sub>GenS</sub>  
the treatment / examination the.GEN patient-GEN / the.GEN doctor-GEN

'the treatment/examination of the patient/the doctor'

c. *Non-quantised change:*

die Sedierung / Ermutigung ✓✓des Patient-en<sub>GenO</sub> / ✓?des Arzt-es<sub>GenS</sub>  
the sedation / encouragement the.GEN patient-GEN / the.GEN doctor-GEN

'the sedation/encouragement of the patient/the doctor'

d. *Quantised change:*

die Heilung / Einlieferung ✓✓des Patient-en<sub>GenO</sub> / \*des Arzt-es<sub>GenS</sub>  
the cure / hospitalisation the.GEN patient-GEN / the.GEN doctor-GEN

'the cure/hospitalisation of the patient/the doctor'

With degrees of affectedness re-operationalised as the continuous AI, the following hypotheses are posited:

1. Analysis of the acceptability ratings should reveal an interaction of 'Linking' ('GenO' and 'GenS' conditions) and 'AI', with the simple slopes of the continuous 'AI' variable exhibiting the following patterns:
  - (a) In the 'GenS' condition, 'AI' should be associated with a negative slope, showing a significant decrease in acceptability with increasing AI, as discussed in section 2.4. Increasing semantic prominence of GenO should thus result in lower acceptability of GenS linking.
  - (b) Following the claim of Ehrich and Rapp (2000) that GenO arguments can always be linked independently of the semantic structure of the underlying base verb, acceptability should be high or at ceiling throughout the range of 'AI' in the 'GenO' condition, without a significant effect of 'AI'. The semantic prominence of GenO should thus not influence acceptability of GenO linking.
2. Acceptability of GenO and GenS linking should be at a similar level at the lower end of the AI range.

#### 2.7.3.4 Participants

Complete ratings were collected from 40 native German speakers with a mean age of 28.3 (range 19–68, 8 male, 32 female). Note that again the data of a number of additional participants were not saved due to technical problems, so their data had to be discarded. For participation in the study, participants could opt for course credits or for taking part in the drawing of a voucher for an online store.

### 2.7.3.5 Data analysis

The binary judgement data were analysed using logistic mixed models (or generalised linear mixed models – GLMMs) via the `glmer` function of package *lme4* (Bates, Maechler, Bolker, & Walker, 2014) for the statistical environment R (R Development Core Team, 2013). Fixed effects included in the model were 'c.AI' and 'c.Linking' as well as their interaction, testing for linear effects only. All variables were mean-centred, as indicated by the prefix 'c.'; note that this reduces collinearity between predictors and changes the interpretation of the intercept estimate to the estimate of the overall mean. 'GenO' was defined as the reference level for the 'Linking' variable. The assessment of significance of fixed effects is performed using *p*-values based upon Wald *z*-scores<sup>10</sup>.

Maximal random effects structures were used for subjects and items, taking a largely design-guided approach as recommended by Barr, Levy, Scheepers, and Tily (2013), resulting in by-subject random intercepts and random slopes for the predictors 'c.AI' and 'c.Linking' as well as their interaction. As grouping factor for the 'Items' random effect 'Story' was used, including by-story random intercepts and random slopes for the 'c.Linking' predictor, which varied within 'Story'. Correlation parameters between random effects were included by default. The final random effects specification for subjects and items in `lmer` notation was thus `(1 + c.Linking * c.AI | Subject) + (1 + c.Linking | Story)`. Random intercept or slope effects for subjects and items were only omitted in the face of convergence problems or indications of overparameterisation of the model (e.g. total correlation between random effects).

Planned post-hoc assessments comprised testing for simple slope effects of 'AI' for each level of 'Linking'. To this end, separate GLMMs were fitted for the 'GenO' and 'GenS' subsets, taking the model specification of the final model for the complete data set as starting point and reducing the fixed and random effects parts as appropriate (i.e. removing any fixed and random effects involving 'Linking').

In addition to the information available by default from the `lmer` output, a number of complementary statistics were computed; univariate 95% confidence intervals for the fixed effects estimates were obtained using the functions `confint` and `glht` of package *multcomp* (Hothorn, Bretz, & Westfall, 2008). Predictions for the continuous 'c.AI' predictor based upon the model's fixed effects estimates were computed for each level of the 'c.Linking' predictor (i.e. corresponding to 'GenO' and 'GenS') using the `ezPredict` function of package *ez* (Lawrence, 2013) along with 95% bootstrap-based prediction intervals using 5000 bootstrap samples. Predictions were computed across the range of the 'c.AI' covariate at 100 equally spaced values between the minimum and maximum value of 'c.AI'. Predictions and corresponding intervals for the difference between 'c.AI' for the 'GenS' and 'GenO' levels were also computed using bootstrap samples.

Model diagnostics included a check of the model fit using the `plotlogistic.fit.fnc` function of package *languageR* (Baayen, 2013), which plots observed proportions against mean

---

<sup>10</sup>Note that the significance of fixed effects was additionally confirmed via likelihood ratio tests using the `anova` function.

predicted probabilities. The influence of outliers was assessed using several functions from package *influence.ME* (Nieuwenhuis, Grotenhuis, & Pelzer, 2012), which allowed to assess the influence of a given story on the final model's estimates refitting it leaving out one story at each iteration using function *influence*. Subsequently, functions *cooks.distance* and *dfbetas* were used to acquire indices of the influence of each story on the model estimates and models were refit using reduced data sets leaving out stories which were above the cut-off thresholds for Cook's distance AND DFBETAS values; the critical thresholds for either value were calculated as suggested by Nieuwenhuis et al. (2012). When the estimates of these reduced models did not significantly differ from those of the full models, the latter are presented without further discussion of possible effects of outlying items.

### 2.7.3.6 Results

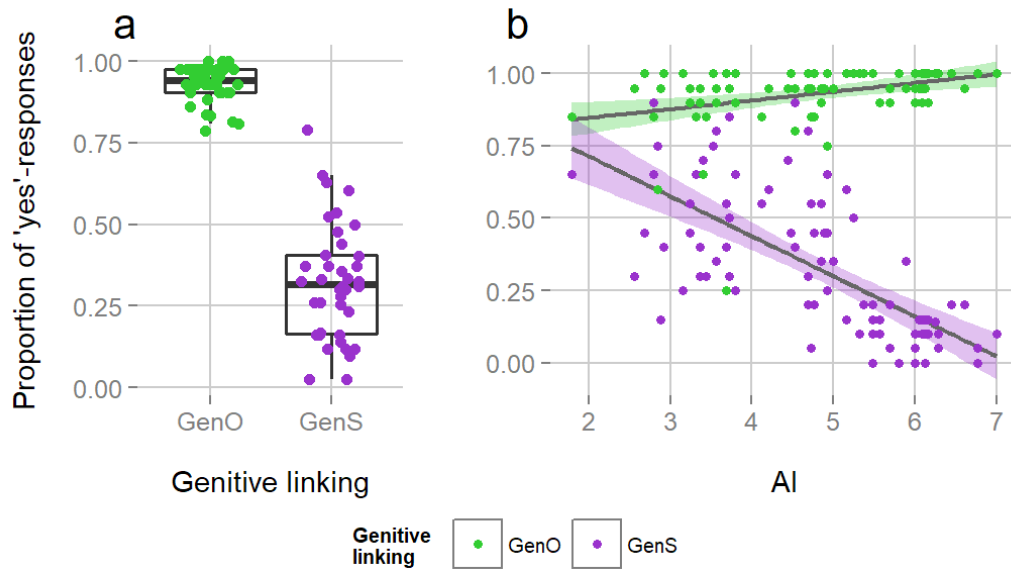
To illustrate the effect of the 'Linking' variable on the acceptability judgements, the proportion of 'Yes' responses for each participant and level of 'Linking' was computed. The boxplots in figure 2.2a show that GenOs were accepted much more often as coherent continuations than GenSs (GenO: mean = 0.93, SD = 0.06; GenS: mean = 0.32, SD = 0.17). Figure 2.2b shows the proportion of 'Yes' responses per nominalisation for each 'Linking' condition over the range of 'AI' as a scatterplot; separate linear smooths for 'AI' are added for the 'GenO' and 'GenS' conditions. The smooth for the 'GenO' condition suggests high acceptability rates throughout the range of 'AI', though acceptability seems to have increased slightly with higher AI values. Conversely, in the 'GenS' condition, acceptability was highest for nominalisations with a low AI and decreases rapidly as AI increases, with the absolute minimum close to zero at the maximum AI values.

The final model specification of the model for the complete data set in *lmer* notation is given in (99) and table 2.1 summarises the random and fixed effects of the model. The upper subtable lists the variances and standard deviations of the included random effects as well as their correlations and the lower subtable contains the following information for each fixed effect: the coefficient estimate (in log-odds) and its associated standard error in the second and third column, 95% CI (confidence intervals) lower and upper bounds in the fourth and fifth columns, followed by the summary of Wald's *z*-test in the next two columns.

```
(99) Yes_Response ~ c.Linkig * c.AI +
      (1 + c.Linkig * c.AI | Subject) + (1 + c.Linkig | Story)
```

The model summarised in table 2.1 shows significant effects of 'c.Linkig' ( $B = -4.87$ ,  $SE = 0.37$ ,  $p < 0.001$ ), 'c.AI' ( $B = -0.2$ ,  $SE = 0.08$ ,  $p = 0.019713$ ) as well as their interaction 'c.Linkig:c.AI' ( $B = -1.61$ ,  $SE = 0.21$ ,  $p < 0.001$ ).

Figure 2.3 summarises the predictions derived from model (99) together with their 95% bootstrap CIs: figure 2.3a shows the predicted slopes of 'AI' for the 'GenO' and 'GenS' conditions in log-odds space and figure 2.3c shows the same transformed to probability space. Both illustrate that the likelihood of accepting a continuation sentence as a coherent continuation



**Figure 2.2:** Proportion of 'Yes' responses for each 'Linking' condition. a: Boxplot of acceptance proportions for each subject and 'Linking' level. Note that the horizontal position of the points varies randomly within a condition to avoid overplotting and enhance visibility. b: Acceptance proportions for each nominalisation at the 'GenS' and 'GenO' levels of 'Linking' plus linear smooths for 'AI'.

was equally high for GenO and GenS occurring with nominalisations associated with a minimal or low AI (estimate at the minimal AI value for GenO: log-odds = 1.85, probability = 0.86; for GenS: log-odds = 1.83; probability = 0.86); as already suggested by figure 2.2a above, acceptability decreases rapidly with increasing AI in the GenS condition to a minimal likelihood at the upper end of the AI scale (estimate at the maximal AI value for GenS: log-odds = -3.37, probability = 0.03), whereas it increases to a maximal likelihood for GenOs (estimate at the maximal AI value for GenO: log-odds = 5, probability = 0.99). Figures 2.3b and 2.3d show the estimates for the corresponding difference between GenS and GenO (i.e. GenS - GenO) across the AI range together with the bootstrap CIs in log-odds and probability space.

Tables 2.2 and 2.3 summarise the post-hoc models fitted to test for simple slope effects of 'c.AI' in each 'Linking' condition, with the model specifications given in (100)<sup>11</sup>.

- (100) a. *GenO-model*:  

$$\text{Yes\_Response} \sim c.AI + (1 \mid \text{Subject}) + (1 \mid \text{Story})$$
b. *GenS-model*:  

$$\text{Yes\_Response} \sim c.AI + (1 + c.AI \mid \text{Subject}) + (1 \mid \text{Story})$$

The model fitted to the GenO subset (table 2.2) shows a significant effect of 'c.AI' ( $B =$

<sup>11</sup>The by-subject random slope for 'c.AI' was removed from the GenO model due to perfect correlation with the random by-subject intercept.

Random effects

Groups	Name	Variance	SD	Corr.			
Story	(Intercept)	0.37	0.61				
	c.Linking	4.26	2.06	-0.19			
Subject	(Intercept)	0.46	0.68				
	c.Linking	2.24	1.50	0.42			
	c.AI	0.01	0.10	-0.63	-0.10		
	c.Linking:c.AI	0.20	0.44	-0.82	-0.26	0.96	

Fixed effects

	Estimate	SE	Lower	Upper	z-value	Pr(> z )	
(Intercept)	1.24	0.17	0.91	1.57	7.40	<0.001	***
c.Linking	-4.87	0.37	-5.60	-4.14	-13.03	<0.001	***
c.AI	-0.20	0.08	-0.36	-0.03	-2.33	0.02	*
c.Linking:c.AI	-1.61	0.21	-2.02	-1.19	-7.64	<0.001	***

**Table 2.1:** Summary tables for the GLMM for the acceptability judgement data.

0.57,  $SE = 0.14$ ,  $p < 0.001$ ) and table 2.3 reveals a significant effect of 'c.AI' in the model fitted to the GenS subset ( $B = -1$ ,  $SE = 0.11$ ,  $p < 0.001$ ). The positive slope coefficient of 0.57 estimated for the GenO model indicates that a unit increase on the AI scale was associated with an increase of the log-odds of a 'Yes'-response of 0.57; conversely, the negative slope coefficient of  $-1$  estimated for the GenS subset indicates a decrease of the log-odds of a 'Yes'-response of 1.

Random effects

Groups	Name	Variance	SD
Story	(Intercept)	1.79	1.34
Subject	(Intercept)	0.53	0.73

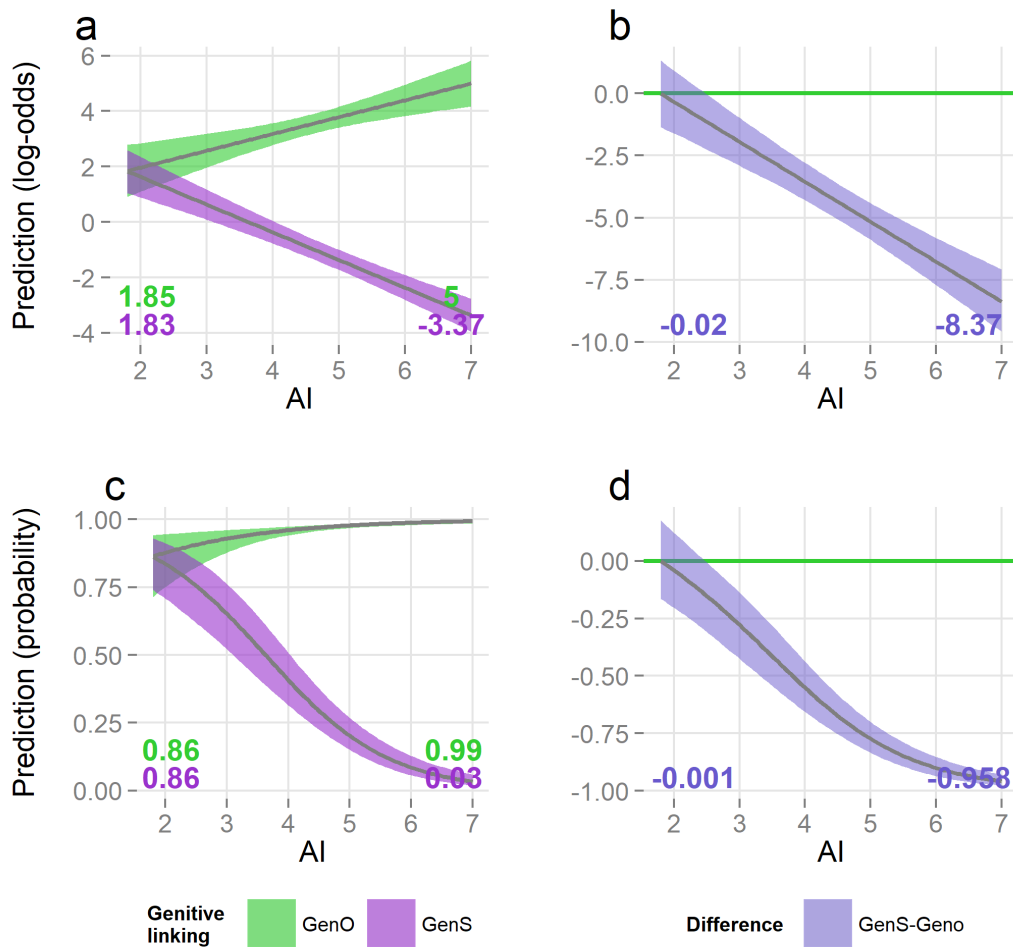
Fixed effects

	Estimate	SE	Lower	Upper	z-value	Pr(> z )	
(Intercept)	3.67	0.27	3.15	4.20	13.64	<0.001	***
c.AI	0.57	0.14	0.30	0.83	4.18	<0.001	***

**Table 2.2:** Summary tables for the GLMM for the GenO subset.

### 2.7.3.7 Discussion

The GLMM analyses of the acceptability judgement data largely confirm the predictions stated in section 2.7.3.3; most noteworthy, the results show a clear interaction of 'AI' and 'Linking', confirming the global prediction of hypothesis 1. The data also confirm hypothesis



**Figure 2.3:** Predictions derived from model (99) with 95% bootstrap CIs. a: Predicted slopes of 'AI' for the 'GenO' and 'GenS' conditions in log-odds space. b: Predictions for GenS - GenO difference across the range of AI in log-odds space. c: Same as figure (a) in probability space. d: Same as figure (b) in probability space. Numbers in the lower left and right corners indicate the estimates at the minimum and maximum values of AI.

1a about the effect 'AI' in the 'GenS' condition: with GenS continuations, an increasing AI is associated with a significant decrease in acceptability, resulting in a sharp drop of acceptance probability from 0.86 at the very low end of the AI scale to 0.03 at its maximum. Note that the acceptance probability for GenS continuations for low AI values is on a par with and does not differ from that of GenO continuations, confirming hypothesis 2. Prediction 1b about the simple slope of 'AI' for 'GenO' continuations is only partially supported by the data: while the estimate generally confirms a high acceptability for GenOs, it also suggests that acceptance probability significantly increases with an increase in acceptance probability to a maximum of

Random effects

Groups	Name	Variance	SD	Corr.
Story Subject	(Intercept)	1.20	1.09	
	(Intercept)	1.45	1.20	
	c.AI	0.10	0.32	-0.56

Fixed effects

	Estimate	SE	Lower	Upper	z-value	Pr(> z )	
(Intercept)	-1.2	0.23	-1.64	-0.75	-5.25	<0.001	***
c.AI	-1.0	0.11	-1.23	-0.78	-8.81	<0.001	***

**Table 2.3:** Summary tables for the GLMM for the GenS subset.

0.99 for nominalisations with a maximum AI. This increase was not predicted by hypothesis 1b.

The results of the study lend support to the hypothesised role of semantic prominence of the GenO argument – measured via its implied degree of affectedness – in the licensing of GenS linking with eventive *-ung*-nominalisations: an increase in the degree of affectedness and thus also in semantic prominence of GenO resulted in decreased acceptability of GenS linking. The high acceptability of GenS linked to low-AI nominalisations, their complete unacceptability when occurring with high-AI nominalisations and the intermediate acceptability with intermediate AI values is in accordance with the predicted patterns as discussed in section 2.4 and illustrated in example (86). The gradual nature of this decrease in acceptability further suggests that the re-operationalisation of degrees of affectedness in form of the continuous AI constitutes a valid approximation to the discrete Affectedness Hierarchy introduced by Beavers (2010, 2011) and is able to capture its semantics-based hierarchical relations in a similar – albeit not identical and somewhat more fuzzy – manner. In addition, it emphasises the need to explain the linking patterns in German *-ung*-nominalisations with an approach which allows to capture more graded judgements than possible within the approach developed by Ehrich and Rapp (2000).

Another finding which is not expected within the approach of Ehrich and Rapp is the facilitatory effect of 'AI' for 'GenO' linking. While acceptability for GenO continuations was at a high level throughout the range of the 'AI' predictor, planned follow-up testing revealed a significant increase in acceptability with increasing AI values. This result is not easily explained within Ehrich and Rapp's approach, since the linking of GenO is in principle taken to be licensed independently of the change of state/affectedness entailed for the base verb's object.

While the hypotheses about the (lack of an) effect of the 'AI' covariate for 'GenO' continuations was based on Ehrich and Rapp (2000) and not met by the data, I believe that the current approach can account for all of the findings, since it emphasises the degree of affectedness as the decisive semantic criterion for argument linking in eventive *-ung*-nominalisations.



Remember that a higher degree of affectedness corresponds to a higher degree of semantic prominence of a verb's direct object, which, in turn, can be understood as increased proximity to the Proto-Patient role (see Beavers, 2011, section 1.5.2.1 in the preceding chapter and section 2.4 above). As previously discussed, a lower degree of affectedness conversely implies that the direct object's semantic distance from the Proto-Patient role increases and – at the same time – that the semantic distance to its subject co-argument decreases.

The notions of semantic prominence and proximity to the Proto-Patient role offer an intuitive way of accounting for the findings of the acceptability study. On this account, the semantic prominence/prototypicality of the base verb's direct object influences the acceptability of GenO linking to a certain degree: while GenOs can generally be linked as postnominal genitives (as witnessed by the high acceptance probabilities throughout the AI range), increased proximity to the Proto-Patient role will still somewhat increase acceptability. To explain the acceptability patterns found with GenS linking, one assumption to be made is that higher semantic prominence correlates with the ability of GenO to 'suppress' its GenS co-argument for linking to the postnominal genitive position; this assumption does not seem ad-hoc nor inconsistent with Ehrich and Rapp's original explanation and it offers a natural way of capturing the idea that GenO and GenS compete for linking to the only available argument slot within the German NP.

The competition process can be conceptualised as the ability of the genitive argument NOT presented in a given sentence (i.e. the 'latent' argument) to intrude upon the argument actually realised and interfere with its linking. When a GenS is presented in a sentence, for example, its acceptability thus mainly depends upon the semantic profile of its GenO co-argument roughly in the following way: when the affectedness and thus semantic prominence of GenO maximal, the implicit GenO co-argument can completely suppress its GenS argument, rendering GenS linking ungrammatical, as indicated by the very low acceptance probability in the high AI range. As affectedness/semantic prominence decreases, however, GenO's ability to intrude upon and suppress GenS is lowered, resulting in higher acceptance probability of GenS with an equal degree of acceptance for minimally affected GenOs.

As pointed out by Beavers (2011) and discussed above in section 2.4, the Affectedness Hierarchy can also be conceived of as classifying verbs in terms of their degree of semantic transitivity, which may provide a slightly alternative perspective on the current findings. On this view, the higher the degree of transitivity of a nominalisation, the more unacceptable becomes GenS linking, which may be a reflex related to the observation of Levin (1999) that objects of CTVs cannot be as readily omitted as those of NCTVs (see section 1.5.2.1 in the previous chapter). However, as previously mentioned these slightly different perspectives are rather complementary in nature, since a change in the degree of affectedness/semantic prominence always entails a corresponding change in the degree of transitivity on this unipolar view. In the current context, these two perspectives thus cannot be distinguished empirically.

One way to investigate the possible source of the patterns found in more detail may be provided by shifting the focus onto the degree of semantic transitivity as defined by the semantic distance between the GenO and GenS co-arguments: on this view, the results may be the consequence of increasing (GenO) and decreasing (GenS) semantic distance

between GenO and GenS and the accompanying shift in the degree of semantic transitivity, a view provided by the bipolar, prototype-based approaches to argument realisation of Blume (2000) and Grimm (2011). An empirical test could be undertaken by acquiring a kind of 'semantic distance index' for each nominalisation's base verb, which does not only incorporate the degree of affectedness of the direct object, but at the same time also a semantically based prototypicality index of the degree of agentivity of the respective subject (which may reflect the extent of control and/or volition the subject has in the event); the sum of the object- and subject-related indices could then be taken as a measure of the semantic distance between the co-arguments and a comparison of the fits of such a bipolar vs. a unipolar model may provide more insight into possible origins of the observed effects.

While an embedding of the current findings within a larger theory of grammar is not an aim I will pursue in this thesis, the work of Sorace (2000) and Legendre (2007) may provide a promising starting point with interesting links to the current study. Sorace investigated the effects of semantic properties of *intransitive* monadic verbs on selection of the perfective auxiliaries *be* and *have*; in some languages, verbs select for either of these, as illustrated in (101) for the German verbs *ankommen/sterben* (*arrive/die*), which select *be*, and *arbeiten/schreien* (*work/scream*), which select *have*.

- (101) a. Der Mann ist angekommen / gestorben.  
           the.NOM man.NOM is arrived / died  
           'The man arrived/died.'
- b. Der Mann hat gearbeitet / geschrien.  
           the.NOM man.NOM has worked / screamed  
           'The man worked/screamed.'

While the sentences (101a) and (101b) appear to be identical on the surface, it has long been argued that the syntactic representations underlying them differ systematically: on this view, the sole argument of verbs like *arrive* or *die* corresponds to the argument linked as direct object of a transitive verb, whereas the single argument of verbs such as *work* or *scream* is equivalent to the argument realised as subject of a transitive verb. The former class of verbs has become known as 'unaccusative' verbs and the latter as 'unergative' verbs (Perlmutter, 1978) and they were originally taken to semantically correlate with patient- and agenthood, respectively. This distinction was thought to be able to explain the differences in auxiliary selection (see the discussion in Sorace, 2000, p. 879). However, as Sorace shows using a number of Western European languages, some verbs show consistent auxiliary selection across and within languages, whereas other verbs are subject to variation in auxiliary selection inter- and intralinguistically. To explain these patterns, she posits the 'Auxiliary Selection Hierarchy' as represented in table 2.4: the hierarchy is based upon the two semantic notions of telic change, which correlates with unaccusativity and the selection of *be* and is maximal for verbs showing the semantic features listed at the top of the table, and agentive unaffffecting process, which correlates with unergativity and the selection of *have* and is maximal for verbs belonging to the verb classes listed at the bottom of the table. Verbs belonging to the extremes of this hierarchy are core instances of unaccusative/unergative verbs and show

Change of location	selects <i>be</i> (least variation)
Change of state	
Continuation of a pre-existing state	
Existence of a state	
Uncontrolled processes	
Controlled processes (motional)	
Controlled processes (non-motional)	selects <i>have</i> (least variation)

**Table 2.4:** Auxiliary Selection Hierarchy of Sorace (2000, p. 863).

consistent auxiliary selection within and across languages, whereas those situated at the mid-levels of the hierarchy are more prone to variable selection of auxiliaries.

Legendre (2007) reformulates the Auxiliary Selection Hierarchy in terms of a number of semantic features, including [inherent displacement], [inhomogeneity], [telicity], [directed change], [state] and [inherent volitionality] and translates it into a universal ranking of mapping constraints as applied within the grammatical framework of Optimality Theory (OT – Prince & Smolensky, [1993]/2004). Legendre (2007, p. 1526) presents the hierarchy of lexicon-syntax mapping constraints in (102), where '1' stands for 'subject' and the constraints are to be interpreted as 'avoid' (see also the discussion in Sorace & Keller, 2005, p. 1511). Since this constitutes a translation of an implicational hierarchy of lexical-semantic properties into a ranking of constraints as applied in OT, it may be taken as a starting point for embedding the implicational Affectedness Hierarchy of Beavers (2010, 2011) into a larger framework of grammar and thus also for providing a somewhat broader perspective on the findings of the current experiment on argument linking in *-ung*-nominalisations, though I will not pursue this suggestion further here.

- (102) \*1/+Displacement >> \*1/+Inhomogeneity >> \*1/+Telicity >> \*1/+Directed change  
>> \*1/+State >> \*1/-Inherent volitionality

In a nutshell, the discussion of the results of the acceptability judgement experiment presented in this section suggests that the notion of degrees of affectedness/semantic prominence implemented via the AI, which can be considered an approximation to the Affectedness Hierarchy developed by Beavers (2010, 2011), can explain the graded acceptability patterns found in a parsimonious manner. Importantly, it also suggests that the experimental paradigm employed – which involves argument linking within nominal rather than verbal structures – provides a valid environment for an empirical test of possible processing correlates underlying predicate-induced semantic determinants of argument linking at the semantics/syntax interface.

## 2.8 Summary

In this chapter, I suggested that use of deverbal *-ung*-nominalisations may represent a shortcut, rather than a detour, for investigating the role of affectedness in argument linking experimentally: since all morphosyntactic cues for judging the acceptability of the linking of GenO/GenS are neutralised, the only way left to resolve linking conflicts with eventive *-ung*-nominalisations is to exploit the predicate-induced semantic property of the affectedness of the GenO argument. While the theory of nominal argument linking presented by Ehrich and Rapp (2000) attributes the (un)acceptability of GenS linking to the presence of a *BECOME* operator in the base verb's LSS, I argued that the Affectedness Hierarchy of Beavers (2010, 2011) provides a slightly alternative account of the observed patterns in terms of affectedness-induced semantic prominence, which further allows to capture more graded intuitions about possible linkings in a straightforward way. The results of the acceptability judgement experiment suggest that the re-operationalisation of Beavers' degrees of affectedness in terms of the continuous AI represents a valid approximation to the original levels of the discrete Affectedness Hierarchy. In addition, the findings cannot be fully captured by Ehrich and Rapp's approach, but receive a natural explanation when viewed as the result of semantic prominence of the base verbs' direct object, assuming that higher semantic prominence goes hand in hand with the ability of GenO to suppress the GenS co-argument in the competition for linking to the postnominal genitive position. A complementary and closely related perspective is gained by interpreting the results as the consequence of varying semantic distance between the GenO and GenS co-arguments with an accompanying shift in the predicate's degree of semantic transitivity. The following chapters make use of the experimental paradigm introduced to investigate behavioural and electrophysiological processing correlates underlying the patterns identified in the acceptability judgement experiment.

## Chapter 3

# Psycholinguistic backgrounds

The previous chapter presented an offline judgement task probing the influence of the AI on the acceptability of linking genitive NPs to deverbal nominalisations; before presenting two experiments probing for underlying effects during online processing in the next chapter, the current chapter reviews the most relevant psycholinguistic backgrounds, thus setting the stage for investigating the main research questions about possible processing effects. These were stated in the Introduction and are repeated in the following:

1. Does affectedness have any consequences for the processing of predicates with verbal semantic structure? If so, can we isolate behavioural and electrophysiological processing correlates of this semantic property on the predicate itself?
2. Does predicate-induced affectedness have any impact on the integration and licensing of NP arguments following the predicate within a sentence? If so, can we identify behavioural and electrophysiological processing correlates of such processes on the respective argument NP?
3. How do the findings fit with current models of sentence processing which address these aspects from related perspectives?

In section 3.1, I briefly outline the central aims of the two experiments conducted to provide a context for the literature review following in section 3.2. The review starts with a discussion of studies which have investigated the processing correlates of lexical-semantic complexity of verbs using a number of different methodologies (section 3.2.1); this section will provide the backgrounds for addressing research question one as formulated above. The remainder of this chapter introduces backgrounds related to research questions two and three: section 3.2.2 summarises studies which have examined the interplay of verbal semantics and syntactic processes during online sentence processing. Section 3.2.3 concentrates on electrophysiological markers which have received canonical interpretations as correlates of semantic-conceptual and syntactic processing and discusses recent findings which call their traditional interpretations into doubt. This section sets the stage for presentation of

the Extended Argument Dependency Model (Bornkessel & Schlesewsky, 2006; Bornkessel-Schlesewsky & Schlesewsky, 2009) in section 3.2.4, a model of sentence processing which focusses on identifying core processes underlying establishing predicate-argument relationships. The two final sections provide brief outlines of further relevant aspects, the role of similarity-based interference in sentence processing (section 3.2.5) and electrophysiological markers implied in the processing of discourse structure (section 3.2.6).

## **3.1 From offline judgements to online processing: aims**

As argued in the previous chapter (see section 2.6), the current experimental nominalisation-based paradigm minimises the influence of morphosyntactic processes in the resolution of possible linking conflicts; at the same time, determining the acceptability of the constructions requires to access semantic features of the respective co-arguments which are determined by the lexical-semantic properties of the base verb, operating on and exploiting this information. The two sentence processing experiments presented in the following chapter mainly pursued two aims: the first was to identify possible processing correlates of lexical-semantic interface properties of predicates by isolating effects of AI on the deverbal the *-ung* nominalisations themselves (research question one listed above). The second goal was to trace relevant processing correlates of the impact of the AI for the integration of postnominal genitive arguments (research question two); this potentially allows to reproduce the results of the offline acceptability judgement task during online sentence processing and provides first information about the underlying dynamics of relevant processes during sentence reading.

The two experiments presented below make use of the self-paced reading (SPR) and event-related potentials (ERP) methodologies; information about processing correlates underlying the two aspects of interest gained with these methods may provide a detailed picture of the nature and time course of the role of affectedness in online argument linking. The primary goal of the SPR study was to test whether processing correlates of either of the two aspects can be found on a behavioural level and, if so, at what point in time they arise during sentence reading (e.g. can effects of AI be found immediately upon processing of the genitive arguments?), while the ERP experiment was carried out to investigate the underlying processes in more detail at the electrophysiological level and to isolate potential sub-processes. Before presenting these two studies in the next chapter, I will review work relevant for either of the two aspects of interest in the following sections.

## **3.2 Psycholinguistic backgrounds**

### **3.2.1 Processing correlates of verbal lexical-semantic complexity**

This section presents an overview of psycholinguistic studies relevant for the first aspect of interest, i.e. processing correlates of lexical-semantic complexity of predicates with verbal semantic structure. While experimental investigations of other facets of lexical items – such

as differences and commonalities between nouns and verbs or the neural representation of conceptual categories underlying nouns (see Kiefer & Pulvermüller, 2012; Vigliocco, Vinson, Druks, Barber, & Cappa, 2011, for recent overviews) – abound, psycholinguistic evidence on processing correlates of syntactically relevant semantic features of verbs (i.e. verbal semantic structure) is still relatively scarce.

While initial experimental studies of the impact of verbal semantic complexity failed to produce confirmatory evidence for structurally complex lexical representations of verbs (e.g. Fodor, Garrett, Walker, & Parkes, 1980; Rayner & Duffy, 1986), a number of more recent investigations have been more successful in providing such indications. Gennari and Poeppel (2003) compared the processing of stative verbs like *deserve/possess/love* and eventive verbs such as *break/discover/build* with a single-word visual lexical decision experiment and an SPR study using whole sentences (see example (103) from Gennari & Poeppel, 2003, p. B30).

- (103) a. The retired musician built his second house from scratch. *event*  
       b. The retired musician loved his second child very much. *state*

In both experiments, eventive verbs took longer to process than stative verbs, which the authors attribute to the higher complexity of the lexical-semantic representations underlying eventive verbs. Assuming a decompositional approach to verbal semantics (see section 1.4.1 for discussion), Gennari and Poeppel argue that accessing eventive verbs requires activating additional semantic structure not present with stative verbs, such as the initial state, resulting state and the change from the former to the latter. The processing of this additional semantic structure accounts for the increase in processing time.

The results of this investigation are supported by a series of studies conducted by McKoon and colleagues (McKoon & Love, 2011; McKoon & Macfarland, 2000, 2002; McKoon & Ratcliff, 2003). McKoon and Macfarland (2002), for example, investigated the processing of internally and externally caused change of state verbs, such as *bloom* and *break*, respectively. Also following a decompositional approach to verbal semantics, they assume that internally caused change of state verbs are associated with a structural template of the form  $x$  (*BECOME IN STATE*), while externally caused change of state verbs are linked with more complex templates of the form  $\alpha$  *CAUSE* ( $x$  (*BECOME IN STATE*)). McKoon and Macfarland's hypothesis that activating the more complex externally caused change of state templates would result in longer processing times than accessing the less complex internal causation templates was confirmed by the results of a series of experiments making use of different tasks such as speeded acceptability judgements, whole-sentence-reading and visual lexical decision (also see McKoon & Macfarland, 2000). In a similar vein, McKoon and Love (2011) investigated possible differences in the processing of (externally caused) change of state verbs such as *break* and surface-contact verbs like *hit*, which they assume to be associated with the less complex event structure template  $x$  (*ACT-contact y*). Again, reading the more complex change of state verbs resulted in longer processing times at the single word (as probed by a lexical decision task) and sentence level (whole-sentence-reading and stop-making-sense tasks).

The authors of the studies summarised so far take these results to attest to the psychological reality of lexical decompositions and event templates: McKoon and Macfarland (2002, p. 2) state that “the parts of verbal meaning underlying syntactic behavior are also assumed to underlie sentence comprehension, and so event templates are expected to have observable consequences for comprehension”, an assumption which is at the core of their ‘meaning through syntax’ (MTS) hypothesis, which attributes a prominent role to verbal semantic structure in sentence parsing (see also McKoon & Ratcliff, 2003). The authors further conclude that the studies’ results challenge accounts of verbal lexical representations in sentence parsing which resort to traditional notions of thematic roles/features and some version of a subcategorisation frame<sup>1</sup> (e.g. Boland, 1997; Ferretti, McRae, & Hatherell, 2001; McRae, Ferretti, & Amyote, 1997; Tanenhaus, Carlson, & Trueswell, 1989; Trueswell, Tanenhaus, & Garnsey, 1994). Gennari and Poeppel (2003, p. B34) argue that “[t]he data also suggest that event structure properties are activated during processing, and that these properties subsume those of thematic roles and argument structure. We believe, therefore, that the processing of event structures, rather than thematic roles per se, operates at the interface between syntactic parsing and semantic interpretation.” This conclusion is in general agreement with the move away from thematic roles lists (and related constructions) and towards alternative approaches to representing verbal semantics in theoretical work on argument linking, as discussed in section 1.4 on predicate decompositions and prototype-based linking theories.

Beyond showing that verbal semantic complexity can incur processing cost, the results of the behavioural studies presented above are relevant for the present investigation for a number of reasons: first, the results of the SPR experiment of Gennari and Poeppel (2003) and the stop-making-sense experiment of McKoon and Love (2011) suggest that such effects occur during the processing of sentences as soon as the relevant predicate is encountered. This allows to formulate hypotheses about where in the sentence stimuli to expect effects of lexical-semantic complexity in the SPR and ERP experiments presented below, i.e. on the deverbal *-ung*-nominalisation and/or the immediately following region.

Second, the results of these two studies also allow to formulate hypotheses about the direction of the effect; since the structurally less complex verb templates map onto lower levels of the Affectedness Hierarchy posited by Beavers (2011) (corresponding to verbs unspecified for change/potential for change) and the more complex ones to higher levels (mostly non-quantised and quantised change predicates – see section 1.5.1), a higher degree of affectedness implies higher processing cost. In terms of the continuous AI used in the present work, an increase in the AI associated with a verb should correlate with an increase in the cost of processing the respective deverbal nominalisation, at least if the underlying effect is roughly linear.

Third, as pointed out by McKoon and Love (2011), these findings may also be interpreted in terms of lexical entailments, thus opening up a direct connection to prototype-based theories of argument realisation, such as that of Dowty (1991). On this view, the increased processing cost associated with change of state verbs is caused by the higher number of

<sup>1</sup>Also see the discussion in Bandecchi and Keane (2013).



Proto-Patient entailments carried by the argument realised as object, making these verbs more difficult to process than verbs which do not entail change. Note that such a perspective provides a direct link to Beavers' Affectedness Hierarchy, since Beavers (2010) points out that degrees of affectedness can be understood as analogues to Dowty's Proto-Patient entailments (see section 1.5.1 for discussion): thus, a higher degree of affectedness implies higher semantic prominence associated with a transitive verb's direct object, which in turn can be interpreted in terms of its semantic prototypicality. This connection provides an additional conceptual underpinning for the present study in deriving hypotheses from previous research.

Note, however, that not all previous work supports the hypothesis that an increase in the AI of a predicate should result in higher processing cost. Manouilidou and de Almeida (2013) conducted an SPR experiment which aimed at isolating processing correlates of verbal core features. One of the comparisons conducted was between verbs which imply a change of state and verbs which do not imply a change, with the former presenting a structurally more complex class of verbs according to Gennari and Poeppel (2003). Change of state verbs did not take longer to process in the study of Manouilidou and de Almeida, however, contradicting the results of the studies outlined above.

Whereas the above studies concentrated on processing correlates of the causal structure underlying different classes of verbs, a number of other investigations focussed on aspectual properties of verbs, which often closely correlate with effects of affectedness, as discussed in section 1.5.2.2. Bonnotte (2008) presents the results of a semantic priming experiment involving two different aspectual classes of verbs, non-durative resultatives like *break* and durative non-resultatives like *play*, which are orthogonal along the two relevant semantic dimensions; the former are characterised by the semantic feature values [- durative, + resultative] and correspond to the class of achievement verbs, while the latter represent activity verbs and are defined as [+ durative, - resultative]. Participants performed two different tasks, judging whether a given verb refers to a durable situation in the durativity task and whether it denotes an event with a clear outcome in the resultativity task. Primes were either from the same or opposing verb class or neutral non-linguistic strings. The priming patterns witnessed in this study using French stimulus material yielded supporting evidence for the relevance of aspectual features in the processing of verbs. Following up on this study, Batiukova, Bertinetto, Lenci, and Zarcone (2014) conducted similar semantic priming experiments with largely identical experimental design and tasks in French and Spanish. The priming patterns again suggested that these aspectual features play a role in the processing of verbs and the authors take their results to show that "they belong to the mental representation of verb meaning" (Batiukova et al., 2014, p. 17).

In a series of behavioural experiments, Coll-Florit and Gennari (2011) tested the effect of the temporal duration of events denoted by verbs. Among these, the results of an SPR study showed that the processing of verbs referring to durative events within a sentence takes longer than that of non-durative verbs, which the authors take to imply that information about the temporal duration of events is accessed as soon as the verb is encountered in a sentence. In another of these investigations, Coll-Florit and Gennari collected subjective ratings about the perceived temporal duration of the events denoted by verbs and found

that these continuous duration ratings correlated with the reading times of verbs presented in previous experiments: the longer the duration of an event denoted by a verb was perceived, the longer its reading time. Importantly for the current thesis, this finding implies that verb-semantic factors operationalised as continuous variables can adequately – or even more appropriately – capture relevant effects.

Beyond this behavioural work, some recent studies investigated effects of lexical-semantic complexity of verbs using electro- and magnetoencephalography (EEG and MEG) and functional magnetic resonance imaging (fMRI). Malaia, Wilbur, and Weber-Fox (2011) probed the effects of telicity in the parsing of written reduced relative clauses and their unreduced counterparts. Thus, the reduced relative clauses in (104) only differ in the telicity of the verb within the relative clauses (*spotted* vs. *chaperoned*).

- (104) a. The actress spotted by the writer left in a hurry. *telic*  
       b. The actress chaperoned by the writer left in a hurry. *atelic*

While the focus of this study lay on interactions of the verbal telicity manipulation with subsequent syntactic reanalysis processes (surfacing on postverbal regions – see the following section), analysis of the verb segment itself suggested that the latency of the early N100 component was slightly longer for telic verbs. The authors speculate that early negativities may reflect the allocation of additional attentional resources at the earliest stage of word processing. Note, however, that Malaia, Wilbur, and Weber-Fox (2009) failed to find effects of telicity at the verb itself, despite the use of similar stimulus material.

Focussing on another word class, Steinhauer, Pancheva, Newman, Gennari, and Ullman (2001) investigated electrophysiological effects of lexical-semantic complexity by comparing two different kinds of *nouns* within sentences, mass and count nouns. Count nouns were associated with a more negative ERP at anterior electrodes than mass nouns between approximately 300 and 600 ms post word onset. Assuming that mass nouns constitute the more basic and unmarked class of nouns and that count nouns feature additional semantic structure, such a left anterior negativity may be interpreted as a correlate of lexical-semantic complexity.

Brennan and Pylkkänen (2010) compared the processing of different types of processes leading to semantically induced complexity. One of these processes is coercion, in which the interpretation of an item is ‘enriched’ to resolve a semantic mismatch with another item in a sentence; in a sentence like *The boy began the book*, for example, *the book* denotes a physical object, rather than an event with temporal structure which one may begin. Thus, interpreting such a sentence requires to enrich its meaning to *The boy began doing something with the book*, such as writing or reading it. Coercion was compared to lexical-semantic complexity, which was manipulated by using either psychological verbs of the subject experiencer (such as *love*) or object experiencer (such as *scare*) type. These two classes of mental verbs differ in that object experiencer verbs carry an entailment about the causation of the mental state which is absent for subject experiencer verbs: a sentence like *Mary scared John* entails that Mary caused John to be scared, but *Mary loves John*, conversely, does not entail that John did something to make Mary love him (Brennan & Pylkkänen, 2010, pp. 784–785). In

an initial SPR experiment, both types of semantic complexity manipulations led to increased reading times on the word immediately following the verb, as compared to a control condition. For a subsequent MEG study using the same sentence material, Brennan and Pykkänen hypothesised that lexical-semantic complexity would correlate with enhanced activity over anterior sensors, based on the results of Steinhauer et al. (2001) and Malaia et al. (2009); however, while coercion sentences led to higher activity in a time window around 300 ms, lexical-semantic complexity did not result in any effects, failing to reproduce the increased processing cost for object experiencers observed in the SPR experiment. Thus, the results of EEG and MEG studies on effects of verbal lexical-semantic complexity are currently still inconclusive.

Using fMRI and a delayed match-to-sample task, Romagno, Rota, Ricciardi, and Pietrini (2012) explored the neural correlates of telicity by comparing the processing of telic and atelic Italian verbs, hypothesising that telic verbs should lead to an increased response in brain areas previously implicated in verb processing, due to the additional entailment of a specific temporal endpoint carried by telic verbs. This prediction was confirmed by higher activation levels in the left posterior middle temporal gyrus (pMTG) for telic verbs during the encoding phase of the task, which Romagno et al. take to “indicate that event knowledge and verb processing in this region are specifically related to the representation of telicity” and that the pMTG may represent “specifically that kind of conceptual information which is relevant to morphosyntax” (p. 70). This conclusion is in principle compatible with the findings of Bedny, Dravida, and Saxe (2014), who aimed at disentangling the effects of grammatical class (nouns vs. verbs) and semantic class (objects vs. events) by comparing verbs and nouns from different conceptual-semantic classes, including nouns denoting entities (e.g. *alligator*) and events (e.g. *hurricane*), with a semantic relatedness judgement task. Crucially, event nouns patterned together with verbs in causing an increase in activity in the left MTG when compared to object nouns, a finding which may be interpreted as indicating a central role of the left MTG for the representation of information about events. Note, though, that it is not clear what kind of event information is implicated in this case (grammatically irrelevant semantic content – i.e. general conceptual knowledge about events – or grammatically relevant semantic structure; see section 1.2), since most of the event nouns were basic, underived nouns denoting simple events (such as *hurricane*), which are generally held to lack complex verbal semantic structure (see Grimshaw, 1990, as well as 2.1).

Kemmerer, Castillo, Talavage, Patterson, and Wiley (2008) aimed at isolating neural substrates of a number of conceptual-semantic components of verbs – including change of state related information –, by comparing five classes of verbs: verbs of running (e.g. *run* and *jog*), speaking (e.g. *shout* and *mumble*), hitting (e.g. *hit* and *poke*), cutting (e.g. *slice* and *hack*) and change of state (e.g. *smash* and *crack*), with the change of state component being associated with verbs of cutting and, by definition, change of state. The analysis of the fMRI data obtained with a semantic similarity judgement task suggested a role of left ventral temporal regions (mainly the fusiform gyrus and the inferior temporal gyrus) in the representation of change of state related information. However, Kemmerer et al. interpret the neural correlates identified as reflecting processing of information related to the root

level of verbal meaning – i.e. semantic content –, rather than grammatically relevant event template information (i.e. semantic structure). Since the left ventral temporal areas associated with the change of state property have been implicated in the processing of object properties like shape, color and texture, they suggest that the activation may “reflect visual re-enactments of the idiosyncratic types of physical transformations that are designated by verbs in the Cutting and Change of State classes” (p. 33). On their view, the representation and/or processing of grammatically relevant aspects of verb meaning may be subserved by portions of left inferior frontal cortex, including parts of Broca’s area; this specific separation of the two aspects of verb meaning (event templates vs. verbal roots; semantic structure vs. semantic content) is termed the ‘two-level theory of verb meaning’ and discussed in more detail by Kemmerer and Gonzalez-Castillo (2010).

Since source localisation of possible processing correlates is not a goal of the ERP study presented below, the findings of these imaging studies are of limited use for deriving concrete hypotheses in the current context. While it is not clear in all cases which aspect(s) of verb meaning the identified neural correlates relate to, they nevertheless suggest that it is possible to track correlates of semantic structure and content on the neural level.

### **3.2.2 Interactions of verbal semantics and syntactic processing**

While there exists a growing body of literature investigating the processing of different verb-related semantic phenomena at the sentence level (such as object coercion and aspectual coercion; see Baggio, Choma, van Lambalgen, & Hagoort, 2010; Paczynski, Jackendoff, & Kuperberg, 2014; Pickering, McElree, Frisson, Chen, & Traxler, 2006; Piñango, Winnick, Ullah, & Zurif, 2006; Pylkkänen, 2008; Pylkkänen & McElree, 2007; Stockall, Husband, & Beretta, 2010, for example), experimental evidence on the nature of the interaction of verbal semantic structure and argument integration processes is rather rare. O’Bryan (2003) tested the hypothesis that aspectual information associated with verbs influences sentence parsing using different experimental paradigms (speaker-change detection experiments, a ‘word maze’ task as well as post-hoc reanalysis of SPR data). Comparing reduced relative clauses alike to those exemplified in (104), her findings suggest that the parsing difficulty arising with such garden-path sentences is reduced when the verb in the relative clauses is telic, rather than atelic, as indicated by reduced garden-path effects in the disambiguating region (the preposition ‘by’) in the word maze experiment, for example. O’Bryan took these results to support the hypothesis that verbal event structure information is accessed and put to use in sentence processing.

Malaia et al. (2009, 2011) present two ERP experiments following up on the behavioural investigation of O’Bryan (for discussion, see Malaia, Wilbur, & Weber-Fox, 2013). In line with the behavioural findings, the ERP data suggest that telicity reduces the parsing difficulty, as the atelic condition was characterised by a more negative N100 component than the telic condition on the disambiguating ‘by’ segment or following regions in the sentence reading experiments, especially over frontal electrodes. Malaia et al. (2011) explains this facilitating effect of telicity by arguing that the event template of telic verbs may activate the syntactic

slot for the Patient role; this kind of telicity-induced structural priming – which is absent with atelic verbs – subsequently eases reanalysis of the initially posited subject-initial main clause analysis to the object-initial reduced relative structure and the associated re-assignment of the Agent/Patient roles. Thus, these findings may provide first hints about possible correlates of processes at the interface of verbal semantic and syntax, which are the focus of the current thesis.

Yet, a closer look at the constructions used in these studies reveals a number of crucial differences to the current paradigm, one of the most important being the contrast in the locality of the relevant processes: since these are restricted to NP-internal operations taking place between the nominalisation and adjacent genitive arguments, they are more local in the present experiments. With (reduced) relative clauses, on the other hand, the crucial processes take place between elements separated by a clausal boundary, are thus non-local and involve structurally much more complex constructions. As argued by Bornkessel and Schlesewsky (2006), processes underlying the comprehension of complex syntactic structures like relative clauses may be subject to fundamentally different regularities than those operative in the parsing of less complex core relations, such as predicate-argument structures within simple clauses (see section 3.2.4 below). It is thus not clear if the findings summarised above provide a basis for deriving hypotheses relevant for the present study's aims.

One study thematically overlapping with the current one is Czipionka (2014): she investigated the interplay of verb class and the NP-inherent feature of animacy in a series of sentence processing experiments by manipulating the animacy of the object NP-fillers and the case assigned by a transitive verb to its object (accusative vs. dative). She found evidence that NP-inherent animacy information can interact with verb class in different processing measures, including reading times and ERPs. In her discussion of the findings, one of the possibilities she considers is that the patterns observed may be caused by the interplay of inherent and what she terms 'derived' semantic properties of arguments (corresponding to what I have been referring to as 'predicate-induced'), since nominative-dative constructions are generally considered less prototypically transitive than nominative-accusative ones. Note though that in that work differences in the prototypicality of the constructions in terms of their degree of transitivity correlated with differences in case assignment, again representing some potentially relevant differences in perspective and paradigm to the current work (see also section 2.6 of the previous chapter for discussion).

### **3.2.3 ERP-correlates of conceptual-semantic and syntactic processing: traditional dichotomies and more recent findings**

Since the EPR experiment presented in the next chapter aimed at identifying electrophysiological indices underlying operations at the interface of (verbal) semantics and syntax, this section presents potentially relevant ERP correlates of processes operative in either linguistic domain as identified by electrophysiological studies on language processing and summarises some recent issues with their traditional interpretations (for a review, see Kutas, Van Petten,

& Kluender, 2006).

Kutas and Hillyard (1980) isolated an ERP component which came to be known as the 'N400', referring to a negative deflection of the ERP which starts at about 200–300 ms post word onset, with a core interval from 300 to 500 ms, and a peak at around 400 ms after onset; N400 effects usually display a centro-parietal distribution. A comparison of conceptually incongruous sentence final words such as in (105a) with congruous counterparts (e.g. *butter* instead of *socks*) yielded a more pronounced (i.e. negative) N400 for incongruous endings.

- (105) a. # He spread the warm bread with **socks**. *incongruous*  
b. He mailed the letter without a **thought**. *unexpected*  
c. The Dutch trains are **white** ... *plausible but false*

Beyond such outright conceptual anomalies, N400 effects within sentences have also been obtained by a number of other manipulations, including congruous but unexpected continuations or sentence endings as in (105b) (Kutas & Hillyard, 1984) as well as violations of world knowledge; thus, since trains in the Netherlands are actually yellow, not white, a sentence like (105c) constitutes a plausible statement which is nevertheless not compatible with the general knowledge of Dutch speakers. Such word-knowledge incompatibilities were shown to cause N400 effects alike to those engendered by conceptual incongruities or unexpectedness (Hagoort, Hald, Bastiaansen, & Petersson, 2004). Berkum, Hagoort, and Brown (1999) showed that N400 effects can further be induced by incongruities with the larger, extra-sentential discourse context. In addition, semantic priming and conceptual relatedness of words within word lists modulate the amplitude of the N400, with conceptually supportive contexts resulting in smaller N400s; further, a number of lexical-conceptual properties such as word class, concreteness or frequency have been shown to influence the N400 response (see Kutas et al., 2006; Lau, Phillips, & Poeppel, 2008, for reviews).

In one dominant view of N400 effects, these are thought to index processes underlying 'semantic integration' of critical words with the prior sentence or discourse context (Brown & Hagoort, 1993; Kutas & Hillyard, 1980): words which are unexpected or incongruent given the prior context are more difficult to integrate and thus result in more pronounced N400s<sup>2</sup>. Thus, for a significant period of time, the prevailing functional interpretation of the N400 was that of an electrophysiological correlate of 'semantic' integrative processing.

Studies investigating ERP correlates of morphosyntactic processes, on the other hand, suggested that these are linked to other ERP components, most prominently anteriorly distributed negativities and posterior positivities. These anterior negativities usually have a distribution with a left-anterior focus and have been found to occur with a number of different morphosyntactic phenomena in sentence processing, including phrase structure violations (e.g. Friederici, Hahne, & Mecklinger, 1996; Friederici, Pfeifer, & Hahne, 1993; Hagoort,

---

<sup>2</sup>Note that an alternative to such a post-access interpretation of N400 effects is the lexical view, on which the N400 reflects the ease or difficulty with which features of the long-term memory representations linked to lexical items are accessed (see Lau et al., 2008, for supportive arguments): predictive contexts pre-activate these features, thus facilitating access to them, which in turn results in attenuated N400 amplitudes.

Brown, & Groothusen, 1993; Neville, Nicol, Barss, Forster, & Garrett, 1991; Osterhout & Holcomb, 1992), agreement violations (e.g. Coulson, King, & Kutas, 1998; Friederici et al., 1993) and the processing of long distance relationships such as filler-gap dependencies in well-formed sentences (e.g. King & Kutas, 1995; Kluender & Kutas, 1993).

Some researchers functionally differentiate between two variants of anterior negativities associated with syntactic processing, based upon their temporal occurrence: following this distinction, the Left Anterior Negativity (LAN) occurs between 300 and 500 ms post onset of the critical word, while the Early Left Anterior Negativity (ELAN) precedes the LAN and arises between 100 and 300 ms. The distinction between these two components provides an essential argument for serial models of sentence processing, in which the ELAN represents an index of the earliest stage of structure building devoted to phrase structure building based upon the syntactic category of the respective word; the later LAN, on the other hand, arises during the subsequent processing stage in which thematic relations are established on basis of morphosyntactic processes such as subject-verb agreement (e.g. Friederici, 2002). The LAN has also been shown to play a role in working memory processes underlying the building of syntactic representations and relationships (e.g. King & Kutas, 1995; Kluender & Kutas, 1993).

These anterior negativities are complemented by posterior positivities, which often follow an (E)LAN as part of a biphasic pattern, but may also occur in isolation. The best known of these is the P600, a late positive-going potential which can be observed in the time range of about 500–800 ms and which often peaks around 600 ms post onset, though sometimes it may surface as an extended shift without an easily identifiable peak. Usually, the distribution of the P600 has its maximum at centro-parietal electrode sites. While the P600 has been observed with a number of different kinds of morphosyntactic violations, such as phrase structure violations (Neville et al., 1991; Osterhout & Holcomb, 1992) or agreement errors as in (106a) from Hagoort et al. (1993), it also arises as an ERP correlate in syntactically complex, but grammatical, sentences which require some sort of syntactic reanalysis. An example is the garden-path sentence in (106a), which requires reanalysis from a simple and preferred active transitive reading (e.g. *The broker persuaded the client to ...*) to the reduced relative clause reading (Osterhout & Holcomb, 1992). The P600 has often been interpreted as an electrophysiological correlate of morphosyntactic processing, whether as a general marker of syntactic integration processes (Kaan, Harris, Gibson, & Holcomb, 2000) or an index of late and controlled syntactic repair and reanalysis processes within serial models of sentence processing (e.g. Friederici, 2002; Friederici et al., 1996).

- (106) a. The spoiled child **throw** the toys on the floor. *agreement violation*  
b. The broker persuaded **to** sell the stock was sent to jail. *syntactic reanalysis*

Adding the anterior negativities as additional markers of syntactic processing, the convenient distinction between electrophysiological indices of 'semantic' (N400) and syntactic processing (ELAN, LAN and P600) began to crumble in the face of a number of unexpected discoveries. Thus, manipulations of syntactic properties which were expected to result in

an enhanced P600, but not to affect the N400, surprisingly modulated N400 amplitudes. Hopf, Bayer, Bader, and Meng (1998), for example, presented sentences with sentence-initial, case-ambiguous plural NPs followed by sentence-final verbs which assign either dative or accusative case, as in example (107a). When compared to accusative-assigning verbs, dative-verbs produced a more pronounced N400, rather than an increase in the P600. In the study of Frisch and Schlesewsky (2001), presentation of two animate NP arguments of a syntactically transitive verb which were both marked for nominative case, as in example (107b), resulted in a biphasic N400-P600 pattern, rather than in a P600 only.

- (107) a. Dirigenten<sub>DAT/ACC</sub> ... kann ein Kritiker ruhig  
conductors.DAT/ACC ... can a.NOM critic.NOM safely  
**umjubeln**<sub>ACC</sub>/**applaudieren**<sub>DAT</sub>.  
celebrate/applaud  
'Critics can safely celebrate/applaud conductors.'
- b. \* Paul fragt sich, welch-er Angler **der** Jäger  
Paul.NOM asks himself which-NOM fisherman.NOM the.NOM hunter.NOM  
gelobt hat.  
praised has  
[uninterpretable]

Similarly, a growing number of studies have found P600-effects in contexts where modulations of the N400 were expected. Kuperberg, Sitnikova, Caplan, and Holcomb (2003) visually presented sentences such as (108a), in which the inanimacy of the subject (*eggs*) clashes with the selectional requirements of the respective verb (*eat*); while these violations are of a conceptual-thematic nature, they failed to evoke an N400 when compared to congruous sentences such as *For breakfast the boys would only eat toast and jam*. Rather, they resulted in an enlarged P600, despite the absence of primarily syntactic ambiguities or violations in these constructions. While initial explanations of this phenomenon still referred to syntactic processes for the resolution of this animacy-based conflict (Hoeks, Stowe, & Doedens, 2004; Kuperberg et al., 2003), other studies like that of Van Herten, Kolk, and Chwilla (2005) suggested that such 'semantic P600s' can also be elicited by manipulations not based on animacy: the sentence in (108b) involves a reversal-based pragmatic plausibility issue, which also resulted in an enhanced P600.

- (108) a. # For breakfast the eggs would only **eat** toast and jam.  
b. # The fox that hunted the poacher stalked through the woods.

In her review of such 'semantic' P600 effects, Kuperberg (2007, p. 27) points out that one aspect all the constructions used in these studies have in common is that they involved "verb-argument semantic violations in which the argument(s) could have plausibly occupied alternative thematic role(s) around the critical verb, had the syntax allowed. Simply put, many of these semantically incongruous sentences appeared to be repairable by ignoring the syntax and changing the thematic roles of the critical verb's arguments". This kind of potentially congruous – but syntactically unlicensed – thematic verb-argument relationship has



been characterised as 'semantic-thematic attraction' (Kim & Osterhout, 2005) or 'semantic-thematic fit' (Kuperberg, Caplan, Sitnikova, Eddy, & Holcomb, 2006).

While a number of additional studies have investigated this phenomenon (e.g. Kim & Osterhout, 2005; Kolk, Chwilla, van Herten, & Oor, 2003; Kuperberg et al., 2006), its implications for the functional interpretation of the P600 (as well as the N400) and the validity of the purported clear and sharp distinction between 'semantic' and syntactic processing and the respective ERP correlates in form of the N400 and P600, respectively, remain unclear and different explanations have been forwarded. Van Herten et al. (2005), for example, interpret the P600 as a reflection of monitoring processes which check the veridicality of a sentence analysis. Kuperberg (2007), on the other hand, argues that the N400 and the P600 subserve two separate, but interacting, processing streams: on this account, the N400 (partially) reflects processing operations within a 'semantic memory-based' stream, which "appears to compute the semantic features, associative relationships and other types of semantic relationships between content words (including verbs and arguments) within a sentence" (p. 36). On contrast, the second, 'combinatorial' processing stream "involves the combination of words through algorithmic mechanisms to build up higher-order meaning" (pp. 36–37), which operates partially in parallel to the 'semantic' stream. This second stream carries out computations based upon morphosyntactic information and 'semantic-thematic' cues such as animacy; 'semantic' P600 effects of the kind reviewed above are thought to be the consequence of inconsistencies between the representations output by either of these two streams.

Note that the persistence with which I have been putting the term 'semantics' in quotation marks within the current section is not a coincidence. As already pointed out briefly in section 1.2, a large portion of the neurocognitive work on word and sentence processing deals with aspects of semantic content, rather than semantic structure; a brief look at the example sentences in (105a) and (108) reveals that much of the work on the N400 as an indicator of 'semantic' processing is no exception, since the incongruities involved are of a conceptual sort and do not tap into semantic processing in the narrow linguistic sense (see also the discussion of example (18) in section 1.2). Pykkänen et al. (2011) present these issues in detail and discuss the implications of a series of MEG studies which aimed at isolating correlates underlying semantic processing proper and compared these to markers of processing difficulties caused by inconsistencies with conceptual and world knowledge (e.g. Brennan & Pykkänen, 2008; Pykkänen, Martin, McElree, & Smart, 2009; Pykkänen & McElree, 2007; Pykkänen, Oliveri, & Smart, 2009).

These studies mostly used sentences involving some form of coercion as stimuli in the 'semantics proper' condition as an instance of enriched composition; one type of coercion is complement coercion, which I briefly discussed in the preceding section 3.2.1 and which is again illustrated in example (109). Pykkänen and McElree (2007) compared coercion sentences (109a) to sentences containing verb-complement relations which are incongruous with conceptual/world knowledge (109b) and to control sentences (109b). While the effect of coercion fell into the N400 time window, source localisation suggested that its generators were located in the ventromedial prefrontal cortex and the respective MEG marker was termed 'anterior midline field' (AMF) by the authors. Conceptual incongruity, on the other

hand, which has often been used to elicit N400 effects in EEG studies, did not affect the AMF, but resulted in modulations of different sources.

- (109) a. The journalist began the **article** after his coffee break. *coercion*  
 b. # The journalist astonished the **article** after his coffee break. *incongruous*  
 c. The journalist wrote the **article** after his coffee break. *control*

Such findings suggest that the processes underlying the establishment of conceptual-associative relations and those driving semantic composition within sentences may be subserved by different brain bases. In addition, it underscores the necessity to differentiate between these two aspects of sentence processing and to be cautious in interpreting N400-related findings for deriving hypotheses about possible ERP correlates of the semantically determined processes which are the focus of the current thesis.

Note, however, that the picture is further complicated by the results of two studies which investigated effects of complement coercion and conceptual incongruity in a single experiment using EEG, rather than MEG. Kuperberg, Choi, Cohn, Paczynski, and Jackendoff (2010) as well as Baggio et al. (2010) made use of the same experimental paradigm applied by Pykkänen and McElree (2007), but failed to find frontal effects – rather, coercion resulted in modulation of the N400, as did the conceptual incongruities. While these differences may stem from differential sensitivities of MEG and EEG to neural sources involved, this discrepancy and the above discussion of issues arising with the functional interpretation of the N400, P600 and other components illustrate some of the challenges involved in isolating and interpreting electrophysiological correlates of processes at the syntax-semantics interface.

In the following section, I will outline a comprehensive model of sentence processing which addresses some of these issues in detail; since it provides a rather detailed account of core processes driving predicate-argument linking in online sentence processing, this model will serve as the main source of deriving hypotheses about possible effects in the ERP experiment presented below.

### 3.2.4 The Extended Argument Dependency Model and the processing of core predicate-argument relations

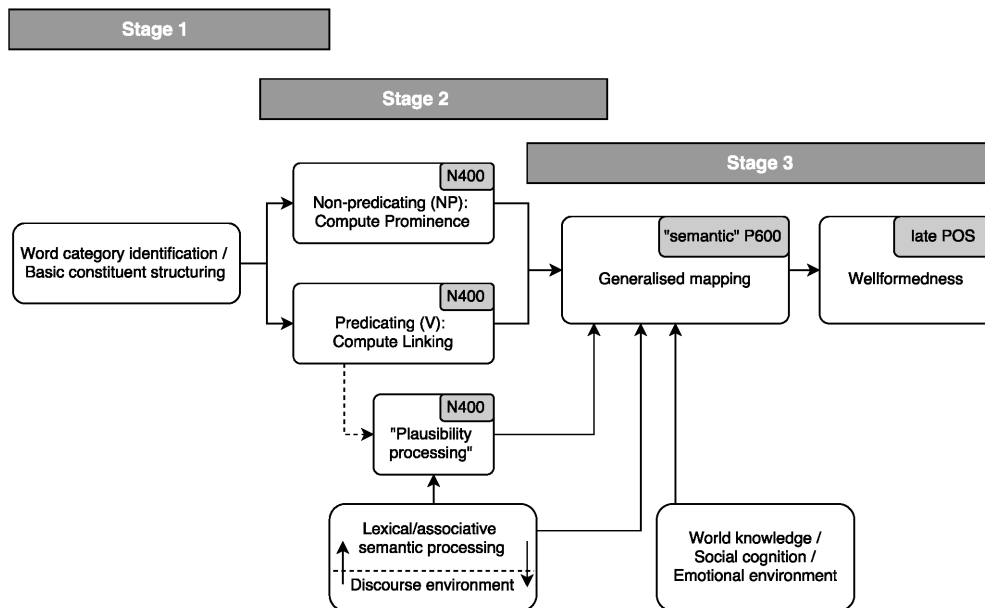
The model I will briefly summarise in the following is the Extended Argument Dependency Model (eADM) of Bornkessel and Schlesewsky (2006); the outline of the model's most important aspects will be based upon the detailed exposition of the eADM provided by Bornkessel and Schlesewsky (2006) as well as relevant additions and modifications presented in subsequent work (Bornkessel-Schlesewsky & Schlesewsky, 2008, 2009; Haupt, Schlesewsky, Roehm, Friederici, & Bornkessel Schlesewsky, 2008, for example). The eADM aims at providing a model of sentence processing at the electrophysiological as well as the neuroanatomical level and posits correlates of its core processing steps at either level; in the following discussion, however, I will focus on the electrophysiological correlates and their functional interpretation within the eADM.

### 3.2.4.1 Scope, architecture and principles

The eADM focusses on the processing of 'core' relations established within a sentence, which includes the relations between a verb (predicate) and its arguments within simple sentences, i.e. single or multi-argument clauses. Non-core relations, such as processing mechanisms behind complex phenomena like reanalysis processes across clause boundaries arising with (reduced) relative clauses or modifier attachment, however, are not at the centre of the model's explanatory domain; in fact, Bornkessel and Schlesewsky (2006, p. 789) assume that these are "subject to fundamentally different regularities to core constituent processing".

The model posits three distinct processing phases (see Bornkessel & Schlesewsky, 2006, pp. 788–805): in phase 1, basic constituent structure is built based upon the word class of the incoming word; this initial structure building is carried out by means of syntactic templates and the resulting representations are devoid of relational information. Phase 2, in contrast, serves to establish these relational aspects, including relations between a verb and its arguments as well as relations between the arguments themselves. A number of different information types are exploited to establish these relations: morpho-syntactic information such as agreement and case markers and positional information, as well as semantic information such as various inherent semantic features of NP arguments. All of these are organised along prominence scales and the resulting prominence information plays a crucial role in establishing the relational aspects of the predicate-argument complex within a sentence. Parallel to and separate from the computation of these relational aspects, plausibility-based processing induced by conceptual-associative relations takes place. These two information types only interact later during phase 3 in a 'generalised mapping' step, in which further information sources such as world knowledge and discourse context are integrated and in which well-formedness and repair processes are triggered when necessary. As Bornkessel-Schlesewsky and Schlesewsky (2009) point out, this multi-stage model of sentence processing belongs to the classes of syntax-first models (e.g. Frazier & Rayner, 1982; Friederici, 2002), since phase 1 serves basic structure-building processes based upon the syntactic category of the incoming word; at the same time, however, the "interpretive burden" (Bornkessel-Schlesewsky & Schlesewsky, 2009, p. 47) imposed on this initial stage and thus also the syntax is reduced by an altered division of labour between the stages. Figure 3.1 provides an overview of the eADM's architecture as presented by Bornkessel-Schlesewsky and Schlesewsky (2009, p. 42).

This reallocation of labour is driven by the 'interface hypothesis of incremental argument interpretation', which underscores the central role of prominence information for the establishment of core relations: "Incremental argument interpretation (i.e. role identification and assessment of role prototypicality) is accomplished by the syntax-semantics interface, that is, with reference to a cross-linguistically defined set of prominence scales and their language-specific weighting" (Bornkessel-Schlesewsky & Schlesewsky, 2009, p. 28). Importantly, syntactic and semantic prominence information are of equal importance, since "prominence features (which would traditionally be classified as semantic or pragmatic) are functionally equivalent to information types such as word order and morphological case marking (which



**Figure 3.1:** Processing phases and architecture of the eADM as presented by Bornkessel-Schlesewsky and Schlesewsky (2009). ERP-markers associated with a given processing step are given in the grey fields in the upper right corners of the respective steps. The figure reproduces figure 4 of Bornkessel-Schlesewsky and Schlesewsky (2009, p. 42).

are traditionally viewed as syntactic)" (Bornkessel-Schlesewsky & Schlesewsky, 2009, p. 20). This is reflected in the types of prominence scales taken to be of relevance, including morphological case marking (e.g. nominative > accusative) and NP-argument order (argument 1 > argument 2) as well as animacy (+animate > -animate), person (1st/2nd person > 3rd person) and definiteness/specificity (+definite/specific > -definite/specific) scales.

Within the eADM, these prominence scales are put to use in answering two central questions arising with the integration of a predicate's arguments during online processing: the first of these is how *role identification* takes place during sentence parsing, i.e. how encountered arguments are assigned a (Proto) role; the second question concerns the influence of *role prototypicality*, i.e. in how far and in what way the semantic prototypicality of NP-arguments in a transitive construction impacts its comprehension. With respect to the first aspect, the eADM differs from other parsing models by assuming that semantic prominence information actually is a factor at all, since the more traditional view considers this to be the domain of the syntax, while role prototypicality is defined via semantic features. Since semantic prominence information is an integral part of the syntax-semantics interface within the eADM, semantics also impacts upon role identification.

A crucial notion within the eADM is the proposal that the semantics-based scales work together to locate an NP-argument of a verb on a scale of 'natural transitivity', which influences

processes related to role identification and role prototypicality. Natural transitivity refers to the idea that certain semantic features of NPs render them more (or less) likely to take on a given function in a transitive event, as with the 'animacy hierarchy' (cf. Comrie, 1989; Lockwood & Macaulay, 2012, see also section 1.3.2, especially example (34)); (see also the discussion of the proposal of semantic transitivity by Hopper & Thompson, 1980, section 1.5.2.1). The eADM implements linking mechanisms developed within prototype-based approaches to argument linking (e.g. Ackerman & Moore, 2001; Blume, 1998, 2000; Davis & Koenig, 2000; Dowty, 1991; Grimm, 2011; Primus, 1999, see also the discussion in section 1.4.2) and follows Van Valin Jr. and LaPolla (1997) in making use of two 'generalised semantic roles' (GR), Actor and Undergoer, which are the counterparts of Proto-Agent and Proto-Patient.

The assignment of these two GRs to the arguments of a transitive verb and the establishment of relational semantic interpretation between NPs are central components of the eADM, with the core processes taking place in phase 2. A core distinction within phase 2 is based upon the dichotomy of predicates vs. arguments: depending upon the syntactic category of the current word identified in phase 1, the parser either enters into 'predicate' processing mode when it is a verb, or into 'non-predicate' (argument) mode when it is a noun/NP, with different processing steps triggered within phase 2 depending upon which of this two streams is entered. The essential mechanisms triggered for argument NPs is 'Compute Prominence', which establishes an interpretive Actor-Undergoer Hierarchy for NPs encountered based upon the morphosyntactic and semantic information organised within the relevant prominence hierarchies. Crucially, this process takes place even when the respective predicating element (usually the verb) has not been encountered yet; in verb-final sentences, which are abundant in German, for example, the prominence information associated with the preverbal NPs is thus exploited as soon as it becomes available. In addition, agreement information available with the NPs is processed by the 'Assign Agreement' operation in phase 2.

The computation of prominence information by means of hierarchically organised prominence information is taken to be a universal, cross-linguistic feature of the sentence parser; higher prominence of an argument NP – whether semantic or morphosyntactic – thus generally favours assignment of the Actor GR to an argument, whereas reduced prominence correlates with assignment of the Undergoer role. The exact role – i.e. the ranking or weighting – of each hierarchy, however, is subject to cross-linguistic variation and thus differs between languages. Bornkessel-Schlesewsky and Schlewsky (2009, pp. 43–44) point out that this cross-linguistic variability also has consequences for the cut-off point for the distinction between role identification and role prototypicality: thus, due to the primacy of argument position for role identification in English, in a sentence like *The cricket ball hit Bill*, the cricket ball can only be interpreted as Actor, despite the fact that it represents a non-prototypical Actor by being inanimate. In other languages, however, the relative ranking of prominence feature types may differ.

A key factor co-determining the ease of processing of transitive constructions is the prototypicality of the co-arguments involved: the more dissimilar these are in terms of their semantic NP-features, the easier the processing. Increased similarity, on the other hand,

causes higher processing costs, since the semantically less distinct GRs will compete with each other. Such an overlap in the semantic profiles of co-arguments can be the consequence of non-prototypical properties in one or either of the co-arguments; according to the animacy hierarchy, prototypical Actors are instantiated by animate and definite NPs, while Undergoers are prototypically inanimate and indefinite. Thus, deviations from these prototypical profiles – such as with inanimate and/or indefinite Actors, for example – decrease the semantic distance between the co-arguments and thus engender increased processing cost. This principle of ‘Distinctness’ which the parser follows holds that “[t]he participants in an event should be as distinct as possible from one another in terms of all available dimensions of prominence” (Bornkessel-Schlesewsky & Schlewsky, 2009, p. 44).

### 3.2.4.2 Relevant ERP-markers and their functional interpretations

The following examples, which are both discussed by Bornkessel-Schlesewsky and Schlewsky (2009, pp. 30–31), serve to illustrate a number of relevant contexts inducing role prototypicality effects and their ERP correlates. Using English stimulus sentences such as illustrated by (110), Weckerly and Kutas (1999) found N400 modulations for animacy-based prototypicality effects. Thus, inanimate head nouns of relative clauses as in (110b) and inanimate relative-clause internal subject arguments as in (110a) modulated the N400 when compared to their respective animate counterparts.

- (110) a. The novelist that the **movie** inspired praised the director for staying true to the complicated ending.  
           b. The **movie** that the novelist praised inspired the director to stay true to the complicated ending.

From the point of view of prototypicality effects, these results can be interpreted as indicating preference violations in both cases, since the inanimate nature of the initial argument NP in (110b) clashes with the expectation that the first NP encountered in an English sentence corresponds to the Actor GR, which is prototypically animate; similarly, the relative-clause internal inanimate argument NP in (110b) is unambiguously analysed as subject/Actor by positional cues, resulting again in an animacy-based conflict.

Reanalysing the original German data of Frisch and Schlewsky (2001), Roehm, Schlewsky, Bornkessel, Frisch, and Haider (2004) found N400 effects for inanimate nominative-marked NPs (111b) which followed initial accusative arguments when compared to animate nominative NPs (111a).

- (111) a. Paul        fragt sich,    welch-en    Angler        der        **Jäger**  
           Paul.NOM asks    himself which-ACC fisherman.ACC the.NOM hunter.NOM  
           gelobt    hat.  
           praised has  
           ‘Paul asks himself, which fisherman the hunter praised.’

- b. Paul fragt sich, welch-en Angler der **Zweig** gestreift  
 Paul.NOM asks himself which-ACC fisherman.ACC the.NOM twig.NOM brushed  
 hat.  
 has  
 'Paul asks himself, which fisherman the twig brushed.'

Within the eADM, such N400 effects are interpreted as markers of increased processing cost caused by prototypicality-related conflicts (Bornkessel & Schlesewsky, 2006; Bornkessel-Schlesewsky & Schlesewsky, 2009) which occur within the Compute Prominence step of the NP/non-predicative processing branch and are *independent* of the verb/predicate (see figure 3.1). This view is related to interpretation of N400 effects arising with case violations such as illustrated by example (107b) above, which Frisch and Schlesewsky (2001) take to index problems of 'thematic hierarchising'. Choudhary, Schlesewsky, Roehm, and Bornkessel-Schlesewsky (2009) obtained N400 effects with incorrect case-assignment in Hindi sentences and argue that the N400 in such contexts is an reflection of misapplications of rules which are 'interpretively relevant'; these stand in contrast to processing problems induced by violations which are purely formal in nature (e.g. agreement violations) and correlate with other ERP markers such as the LAN.

While Compute Prominence is the central processing step within the non-predicate/NP stream of phase 2, the key operation within the predicate/verb branch is 'Compute Linking'. When a verb is processed, this step serves to project the arguments NPs which have already been encountered to the lexical semantic structure of the predicate, which is assumed to be represented in form of a logical structure (LS) via hierarchical lexical decompositions (see section 1.4.1 for an outline and discussion). The NP-related prominence relations which have already been established prior to appearance of the verb are thus mapped onto the hierarchical structure associated with the verb's lexical semantic structure; in addition, other available information sources are drawn upon in the Compute Linking step, including agreement and voice (active vs. passive) markers. When the prominence relations derived from previously encountered NP arguments align with the hierarchical relations encoded by the verb's LS, linking proceeds smoothly, while processing costs ensue "whenever the assumptions about role identification and role prototypicality made prior to the verb do not straightforwardly map onto the verb's LS" (Bornkessel-Schlesewsky & Schlesewsky, 2009, p. 46). Note that in cases in which no arguments have been processed prior to the verb, Compute Linking makes predictions about following NP-prominence relations based on the predicate's LS.

In the detailed exposition of the model provided by Bornkessel and Schlesewsky (2006), increased costs within the Compute Linking step may be associated with a number of ERP correlates, including early parietal positivities, the LAN, N400 and P600. Which (combinations) of these are predicted to arise as a consequence of increased cost in Compute Linking depends upon the underlying source of the problem, resulting in a complex set of eliciting conditions and contexts. An example of an N400 elicited by a linking conflict occurring on the clause-final verb in German sentences is given by (112): here, the argument NPs preceding the verb are both ambiguous between nominative, accusative and dative case and the ambiguity is only resolved by the number marking on the verb, which establishes an agreement

relation with the second NP, which thus receives nominative case, rather than accusative case as predicted by Compute Prominence. The first NP *Peter* is reassigned dative due to the lexical properties of the verb, resulting in an N400 on the predicate (Bornkessel, McElree, Schlesewsky, & Friederici, 2004).<sup>3</sup>

- (112) ... dass Peter Studentin-nen  
 ... that Peter.NOM/ACC/DAT.SG students.NOM/ACC/DAT.PL  
**folgt-en.**  
 followed-PL<sub>→NOM/DAT</sub>  
 '... that students followed Peter.'

In the more recent discussion of the eADM presented by Bornkessel-Schlesewsky and Schlesewsky (2009), the N400 is highlighted as the most prominent correlate of Compute Linking (see their figure 4 on page 42, for example, reproduced here as figure 3.1). The interpretation of such predicate-related N400 effects as a 'reanalysis N400' receives support from (Haupt et al., 2008), who argue that the N400 is the true correlate of 'local' function reanalysis, i.e. subject-object reanalysis (as well as object-object reanalysis, see Hopf et al., 1998) processes which take place *within one and the same clause*: the "reanalysis N400s thus appear to depend on interpretively relevant, but structure independent *grammatical* properties. We would therefore expect to observe N400 effects of this type whenever different prominence-based information sources provide conflicting information about the argument hierarchy within a single clause" (Haupt et al., 2008, p. 87; emphasis in the original). Thus, within the eADM the N400 emerges as a key correlate of linking-related processes at the syntax-semantics interface, which occurs with NP-related prominence conflicts (in the Compute Prominence step) as well as with mismatches in the process of mapping arguments to the lexical-semantic structure of the predicate (Compute Linking).

Note that this functional interpretation of the N400 is not straightforwardly reconcilable with its standard interpretation as an index of conceptual-integrative processing given the prior context (see section 3.2.3 above). Indeed, the eADM distinguishes between two types of N400s: those which are triggered by core processes and a 'plausibility N400', which corresponds to the traditional notion of the N400 as reflecting lexical-conceptual integrative processing related to the processing of non-core aspects. This one-to-many form-to-function mapping is a rather novel and non-standard interpretation of N400 effects which has received support from studies which have found different EEG frequency profiles for core and non-core associated N400s (Roehm, Bornkessel-Schlesewsky, & Schlesewsky, 2007; Roehm et al., 2004). While these functionally distinct N400-types are all part of phase 2 and are thus assumed to be generated in parallel, the underlying core and non-core related processes take place separately and do not interact until the subsequent phase 3, in which the outcomes of both are mapped onto each other (see figure 3.1).

<sup>3</sup>Note that Bornkessel and Schlesewsky (2006) argue that superficially identical function reanalysis involving an nominative/accusative case assigning verb, rather than a nominative/dative predicate, results in a P600 only, rather than an N400 (Bornkessel et al., 2004). Subsequent research, however, showed that function reanalysis involving nominative/accusative verbs also yields an N400 (Haupt et al., 2008).



Within phase 3, this mapping is accomplished by the Generalised Mapping step, which is indexed by a late positivity. Bornkessel-Schlesewsky and Schlesewsky (2008) argue that the 'semantic P600' effects discussed in the previous section (see Kim & Osterhout, 2005; Kolk et al., 2003; Kuperberg, 2007; Kuperberg et al., 2006, for example) correspond to this late positivity and are a consequence of incompatibilities of core and non-core associated processing output of phase 2. Similar to the one-to-many mapping of the N400 in phase 2 of the model, the eADM posits more than one type of late positivity associated with phase 3, each of which has a different functional interpretation. In addition to the late positivity arising with the Generalised Mapping step, a late positivity may also index operations driving a well-formedness check, which monitors for and identifies ill-formed or marked structures. Haupt et al. (2008) maintain that such a well-formedness related late positivity occurs as part of a biphasic N400-late positivity pattern. According to Haupt et al., a monophasic late positivity, on the other hand, indexes reanalysis processes for complex structures with cross-clausal dependencies, such as arising with relative clauses; in the eADM, it is only these reanalysis processes crossing clause boundaries which also involve a phrase structural reanalysis (while clause-internal operations do not entail phrase structure reanalysis processes) and in the terminology of Haupt et al., associated late monophasic positivities qualify as a 'P600' proper. Bornkessel and Schlesewsky (2006) point out the need for further investigation into this functional distinction between late positivities and a P600 (see p. 804 for discussion).

### **3.2.4.3 The eADM and the present work: common ground and differences in perspective**

The eADM is based upon a number of assumptions which allow to establish immediate links to the topic of this thesis. The first is the model's focus on the interpretation of arguments and the processing of local, clause-internal predicate/argument linking operations with structurally simple constructions. Other crucial connections are provided by the eADM's interface hypothesis, which allocates a central role to the semantic prominence of NP-arguments for these processes and the related notion of 'natural transitivity' (for work involving interactions of animacy and verb type during sentence processing, see Czypionka, 2014, for example). The interface hypothesis also underscores the general importance of semantic prominence information as a grammatically relevant type of information active at the syntax-semantics interface. Further, the model's utilisation of prototype-based approaches to argument linking motivates the role of semantic similarity of NP co-arguments and thus the principle of Distinctness.

As discussed in detail in the preceding two chapters, all of these issues also feature prominently in current theories of the role of affectedness in argument linking as well as in the present experimental investigations of semantically determined processes underlying linking patterns in German deverbal *-ung*-nominalisations. One of the theoretical motivations for the current work is provided by the Affectedness Hierarchy of Beavers (2011), which views effects of affectedness largely as following from semantic prominence of a verb's argument, where stronger truth conditions about the degree of change implied for an object correl-

ates with higher semantic prominence. Further, in his discussion of the Affectedness Hierarchy for processes determining argument linking patterns, Beavers (2010) couches this approach to affectedness within prototype-based theories of argument linking (especially that of Dowty, 1991, see section 1.5 for in-depth discussion). In the discussion of the results of the acceptability judgement study in the previous chapter (section 2.7.3.7), I argued that the patterns witnessed can receive a natural interpretation when viewed as a consequence of affectedness-induced semantic prominence of the argument realised as direct object; I further pointed out that on a complementary and closely related perspective, the findings can be interpreted as the consequence of in-/decreasing semantic distance between the GenO and GenS co-arguments and a related shift in the predicate's degree of semantic transitivity. Thus, the eADM provides numerous model-theoretic starting points for deriving hypotheses about possible effects in the ERP experiment presented below. A prime candidate for an ERP correlate involved in the resolution of affectedness-related linking conflicts is the N400, since it has emerged as the central ERP marker of processes underlying prominence-based conflicts within the argument stream of the eADM and linking conflicts within its predicate branch.

Nevertheless, at second glance, there are a number of significant differences between the eADM's emphases and those of the current paradigm, which largely stem from the fact that the prominence of argument NPs is approached from diverging points of departure: in the eADM, the semantic aspects of argument-related prominence are either inherent to the head noun of the respective argument NP (as in the case of animacy) or assigned locally within the syntactic domain of the head noun. In either case, these features are determined independently of the verbal predicate. The focus of this thesis, on the other hand, lies on processes which are triggered by properties of NP-arguments which are induced by the semantic profile of the *predicate*, rather than by NP-inherent characteristics.

This difference in the origins of the relevant prominence hierarchies has consequences in several respects: conflicts involving NP-inherent semantic prominence relations between co-arguments are handled by operating on information which is verb-independent and can be read off argument NPs in isolation<sup>4</sup>. In terms of the architecture of the eADM (see figure 3.1), this may be interpreted as implying that all relevant processing operations take place within the argument/NP stream of phase 2. The processes which are the focus of the present work, on the other hand, require accessing semantic properties of NPs which originate within the verb/predicate stream of phase 2, i.e. in a way they involve operations cross-cutting the two processing streams of phase 2. This issue is related to the category of the word at which effects of prominence and linking conflicts are respectively measured in the studies the eADM refers to, with correlates of the Compute Prominence operation assessed at NP-arguments and those of Compute Linking – which refers to the lexical-semantic representations of predicates – at verbs. One of the current study's aims, however, is to isolate processing correlates underlying the interplay of verbal semantics and linking processes by

---

<sup>4</sup>Note that this is also true for verb-dependent grammatical prominence information on argument NPs such as morphological case marking, which also contributes to argument prominence in the eADM, as outlined above. While case marking on argument NPs is determined by the respective verb, it can be identified on the NPs in isolation of the verb.

identifying integration processes which depend on verb-induced semantic properties of NPs at the *argument*, rather than the predicate.

Further, as discussed in section 1.3.2, Levin and Rappaport Hovav (2005) caution against confounding the effects of NP-inherent properties organised in natural prominence scales with the effects of thematic hierarchies in argument realisation phenomena, since the feature values of NP-arguments drawn from these different prominence scales tend to correlate in position across scales; thus, “[p]roperties of the NPs that typically fill certain semantic roles, which are often confused with properties of the semantic roles themselves, may affect the surface realization of these NPs, but this surface realization most probably does not reflect true thematic hierarchy effects” (Levin & Rappaport Hovav, 2005, p. 183). Bornkessel-Schlesewsky and Schlesewsky (2009, p. 49) acknowledge that “it remains to be clarified how many additional prominence dimensions need to be assumed” in addition, for example, to the animacy and definiteness scales and how these various prominence scales interact. The present work can thus be seen as an attempt to contribute to this issue by investigating the role of predicate-induced semantic prominence of NP-arguments and associated processing correlates. The distinctions between argument-inherent and predicate-induced prominence relations pointed out may give reason to assume that processing conflicts related to either of these correlate with different processing markers.

Other relevant differences are the consequence of the specific experimental paradigm applied in this thesis, which makes use of linking phenomena within the nominal domain. While the core relation between the eventive nominalisation and the following genitive argument (GenS or GenO) is local by being internal to the ‘extended projection’ of the head noun (in the sense of Grimshaw, 2005a) without crossing any clause boundaries and is thus within the scope of the eADM, there are nevertheless additional factors which come into play and may qualify the local nature of the core-relations involved. Remember that the current experimental design is such that the target linking configuration (i.e. which argument is realised as subject and object) is determined in a first context sentence and subsequently used as a basis for estimating the compatibility of a given nominalisation-genitive combination with this established argument linking in the second (continuation) sentence. Thus, the relevant standard of comparison for any predicate-argument sequence in the second sentence is provided by the discourse representation established on basis of the first sentence, i.e. a non-local information source. With the current paradigm, there is thus an interplay of local processes with sentence-external information sources and it is not entirely clear to me how to reconcile this with the role of the N400 as a marker of prominence-based conflict processing within local constructions, for example. In section 3.2.6 below, however, I will present some findings from studies on discourse processing and will discuss some relevant points.

While these are among the most immediate issues arising with the attempt to base predictions about relevant effects in the following ERP-experiment on the eADM, they are probably not the only ones that come to mind (for example other factors such as the exact nature of the task involved, which is usually about establishing an argument linking in the first place versus checking for consistency with a given linking in the current experiments) – nevertheless, I believe that there is enough common conceptual ground to use the model as a well-grounded

point of departure and context for discussion. A more detailed discussion of predictions derived from the eADM about the interplay of affectedness and argument linking processes in the current paradigm are provided below in section 4.2.4.

### 3.2.5 Similarity-based interference in sentence processing

Bornkessel-Schlesewsky and Schlewsky (2009) point out that the principle of Distinctness is compatible with approaches which highlight the impact of interference based on the similarity of certain sentence elements in language processing. Gordon, Hendrick, and Johnson (2001) investigated the effect of the similarity of co-argument NPs in a series of SPR experiments of object- and subject-extracted relative clauses (and clefted sentences). In this study, similarity was manipulated by using different types of NPs, such as descriptions (e.g. *the lawyer*) or names (e.g. *Joe*). (113a) provides an example of an object-extracted relative clause with a description and a name and (113b) illustrates the subject-extracted relative clause counterpart.

- (113) a. The barber that the lawyer/Joe admired climbed the mountain.  
b. The barber that admired the lawyer/Joe climbed the mountain.

Gordon et al. showed that the type of the NP modulates the difficulty associated with the processing of object-extracted relative clauses on the embedded verb and the following region: when the NP-type differed between the two NPs (e.g. when one is a description and the other one a name), the processing difficulty associated with object relative clauses was substantially reduced or even eliminated. The authors attribute this effect of NP-type to similarity-based interference between the NPs; thus, human memory is susceptible to interference when encoding, storing or retrieving items and “[w]hen the NPs are drawn from the same class they are more similar and hence can interfere more with each other’s processing” (Gordon et al., 2001, p. 1420).

Van Dyke and McElree (2006) showed that such similarity-induced interference can not only be triggered by parallel NP-types, but also by lexical-conceptual relatedness between a verb and nouns. In their SPR experiment on clefted sentences, participants were instructed to memorise lists of three nouns presented prior to reading an experimental sentence. Interference was manipulated via the plausibility of the memorised nouns acting as objects of the critical verb: with a memory list of *table/sink/truck*, for example, the memorised words are all plausible objects of the verb *fix* in example (114). Thus, when the verb *fix* is encountered, the memorised items interfered with *the boat* in being linked as the verb’s direct object and processing difficulty ensues, as implied by increased reading times on the verb. With a verb such as *sail*, on the other hand, no such interference arose.

- (114) It was the boat that the guy who lived by the sea sailed/fixed in two sunny days.

The results show that similarity-based processing difficulty is not only caused by the similarity of the type of NPs, but may be related to “linguistic properties of the head verb, including

subcategorization and thematic role information, and experience-based information associated with the verb's lexical entry" (Van Dyke & McElree, 2006, p. 164), which may cause problems at the stage of retrieval. The study thus illustrates that the sources of interference include lexical-semantic properties of items. This is in line with the assumption underlying the principle of Distinctness that overlapping features of co-arguments can cause difficulties in sentence processing.

Leiken and Pykkänen (2014) conducted an MEG experiment on the time course and neural generators of similarity-based interference in sentence processing, where similarity was defined via the type of the NPs, similar to the study of Gordon et al.. Minimum norm analyses of the MEG data suggested a prominent role of the left inferior frontal gyrus (LIFG) in the processing of similarity-based retrieval interference, with an early significant time window peaking at around 250-350 ms and a later one at around 600-700 ms.

The findings of Leiken and Pykkänen's MEG study exhibit interesting parallels to an fMRI study conducted by Hindy et al. (2012), who investigated the neural correlates of representing the change of state an object undergoes in the course of an event in two sentence reading experiments. Example (115) illustrates the paradigm used in one of the two experiments: here, depending on the lexical filler of the direct object of the verb *stomp on* in the first sentence, either no/minimal change is implied for the object (when it is *penny*, for example) or the object is implied to undergo substantial change (*egg*).

(115) The girl will stomp on the penny/egg. And then, she will look down at the penny/egg.

The authors hypothesised that reading about an object which undergoes a change of state results in a conflict caused by the *dissimilarity* caused by the incompatibility of the object NP's (intact) initial state and its resultant state after the event: they argue that such constructions involve maintaining multiple instantiations of one and the same object in different states. Since an object cannot co-exist in two different states at any one point in time, these distinct representations compete, and retrieving the representation of either state involves suppressing the other state, which comes at a cognitive cost. The dissimilarity of the pre- and post-event object states was operationalised by a continuous change index obtained on basis of subjects' ratings. The results again suggest that left inferior frontal regions play an important role in the resolution of this change-related dissimilarity conflict, since the rated degree of change parametrically modulated these brain areas. In addition, activation within this region overlapped with that witnessed for conflict trials in a Stroop colour-word interference task, which Hindy et al. take to imply that the processing of the change-induced dissimilarity conflict is subserved by processes also involved in general conflict resolution as well as conflicts caused by incompatible semantic representations (Thompson-Schill, Bedny, & Goldberg, 2005; Thompson-Schill & Botvinick, 2006). Thus, the results of the studies of Hindy et al. and Leiken and Pykkänen point to a central function of left inferior frontal regions in the processing of (dis-)similarity-based conflicts in language processing.

From the perspective of the present work, the findings discussed above establish a link between the eADM's principle of Distinctness – and thus also prominence-based conflicts in

sentence processing – and current research on the role of similarity in language processing, thus providing a further context for discussing effects of verb-induced prominence/similarity of arguments in the processing experiments on argument linking presented below. While identification of brain areas underlying possible effects is not the aim of this thesis, the agreement on the neural substrates associated with NP-type induced similarity conflicts and change-related dissimilarity competition implied by the studies of Leiken and Pykkänen (2014) and Hindy et al. (2012) suggests a key role of (left inferior) frontal brain regions in the resolution of such conflicts. In addition, the linear correlation between the change-ratings and activation in left inferior frontal cortex in Hindy et al. provide a first hint that change of state induced (dis-)similarity related processing conflicts can be modelled by continuous co-variables such as the AI used in the present study.

### 3.2.6 Relevant aspects of discourse processing

In this final section, I will briefly present a number of findings from ERP research on discourse processing. Since the current experimental paradigm involves inter-sentential processes, these may be relevant to interpret results of the following ERP experiment, especially in the context of the eADM, as discussed above. This brief outline is mainly based on Van Berkum (2012), who provides a review of discourse-related ERP correlates.

One ERP component reflecting aspects of discourse processing is the N400, which may – inter alia – be triggered by words which are incongruous or unexpected given prior extra-sentential information. In the following example (116) from Berkum et al. (1999), when the final word of the third sentence is *slow*, rather than *quick*, it is at odds with the preceding discourse context provided by the first two sentences and elicits an N400.

- (116) As agreed upon, Jane was to wake her sister and her brother at five o'clock in the morning. But the sister had already washed herself, and the brother had even got dressed.  
Jane told the brother that he was exceptionally **quick/slow**.

As Van Berkum (2012) discusses, such discourse-induced N400 effects can receive an interpretation within one of the standard approaches towards N400 phenomena within single sentences, i.e. they may reflect processes of conceptual integration or, alternatively, long-term memory retrieval. Within the eADM, such discourse-related N400 effects are treated as instances of the plausibility N400, which occurs in phase 2 in parallel to – but independent of – N400 effects indexing processes of prominence computation or local grammatical function reanalysis.

Other studies have found P600/late positivity effects with discourse-based processing difficulty. Burkhardt (2006) conducted an experiment which aimed at identifying ERP correlates of updating a reader's mental model of the discourse situation, i.e. discourse memory. She manipulated the complexity of discourse integration by presenting different initial context sentences (117a)–(117c) before a second sentence which contained a critical NP (*the conductor* in (117d)) was presented. The ease of integration of *the conductor* in the target

sentence depends upon the context sentence: with an initial sentence like (117b), the conductor is explicitly introduced into the discourse and thus represents a given NP in the target sentence, while (117c) provides no cues whatsoever about an upcoming conductor, which is thus a new discourse entity in the second sentence. (117a), on the other hand, does not explicitly introduce the conductor, but provides a discourse context (a concert) in which a conductor plays a salient role; consequently, while the NP *the conductor* is not an established discourse entity when it is encountered, it can be integrated into the discourse model by means of inferential 'bridging' processes.

- |       |  |                           |
|-------|--|---------------------------|
| (117) | a. Tobias visited a concert in Berlin.             | <i>bridged NP context</i> |
|       | b. Tobias visited a conductor in Berlin.           | <i>given NP context</i>   |
|       | c. Tobias talked to Nina.                          | <i>new NP context</i>     |
|       | d. He said that the conductor was very impressive. | <i>target sentence</i>    |

Results showed that entirely new NPs correlated with an enhanced N400 and P600 when compared to given NPs; bridged NPs elicited a stronger P600 than given NPs. The topography of this late positivity which occurred in these conditions had a left posterior focus, rather than a more standard posterior midline centre. Burkhardt (2006) interpreted these late positivity effect as reflecting the cost of reorganising or updating the current discourse model (see also Burkhardt, 2007).

Another ERP component associated with the processing of discourse structure is the referentially induced frontal negativity (Nref) identified by Van Berkum, Brown, and Hagoort (1999). They presented short stories such as that presented in (118) (from Van Berkum, 2012), where an NP in the final target sentence (here *the lecturer*) was either referentially unambiguous (when *the lecturer* as well as *the professor* were introduced in the preceding sentences) or referentially ambiguous (when two lecturers were mentioned prior to the final sentence).

- (118) In dismay, the faculty dean called the lecturer and the professor (the two lecturers) to his office. This was because the lecturer (one of the lecturers) had committed plagiarism, and the professor (the other one) had faked some of his research data. The dean told **the lecturer** that there was ample reason to sack him.

When compared to referentially unambiguous NPs, ambiguous ones elicited a sustained negativity over frontal electrodes, which emerged as soon as 300 ms. Subsequent research confirmed that this component termed Nref is associated with processing referential ambiguity with full NPs as well as with pronouns, in both written and spoken language processing (for reviews and discussion, see Van Berkum, 2012; Van Berkum, Koornneef, Otten, & Nieuwland, 2007). Van Berkum et al. (2007) points out that this process of reference establishment is rather fast and views the Nref as reflecting processes of checking the discourse model for suitable entities; Van Berkum (2012) lists intensive inferential inferencing as well as higher load on frontal working memory systems as plausible candidates for functional sources of the Nref.

In this brief outline, I have presented three ERP correlates of discourse-related processing indexing distinct operations. Two of these – the discourse-linked P600/late positivity and the Nref – at least partially reflect a search, and possibly an update, of the discourse model to accommodate currently processed event participants; since such a check or attempted update of the situation model is also part of the experimental task applied in the following ERP experiment, these two components can be taken as candidates for electrophysiological correlates. In the following chapter, I will return to this issue in the discussion of the results of the ERP experiment.

### 3.3 Summary

This chapter presented the relevant psycholinguistic backgrounds and set the following two online processing experiments into context. Two foci were the behavioural, ERP- as well as fMRI-correlates of verbal lexical-semantic complexity (relevant for addressing research question one) as well as processes underlying the interplay of verbal semantic factors and syntactic processes (research question two), even though data on the latter are currently relatively scarce. The review of the eADM (Bornkessel & Schlesewsky, 2006; Bornkessel-Schlesewsky & Schlesewsky, 2009) revealed a number of relevant thematic intersections with the current topic and aims, especially the model's stress on identifying operations driving the establishment of core predicate-argument dependencies in terms of hierarchical prominence relations, including semantics-based prominence features of argument NPs. The eADM also makes use of prototype-based principles of sentence processing, as reflected in its principle of Distinctness, for example; these aspects of the model provide valuable links to the current investigation, even though I have identified and discussed a number of substantial differences. These largely stem from the eADM's focus on NP-inherent prominence features, while the kind of semantic prominence at issue in the current work is attributed to an argument NP by its predicate. The chapter provided a discussion of recent work on the role of similarity-based interference in sentence processing, which is in agreement with the eADM's Distinctness principle and provides an interesting perspective for viewing possible processing effects in the subsequent experiments as arising as a consequence of semantic similarity between co-arguments. The review closed with a brief outline of key results of ERP research on discourse processing. In the following chapters, the findings and principles discussed will serve to derive predictions for the two processing experiments and to interpret possible processing correlates identified.



## Chapter 4

# Affectedness as a factor in online argument linking: reading time and electrophysiological evidence

This chapter presents the two online sentence processing experiments investigating the behavioural and electrophysiological correlates of activating lexical-semantic information about the AI associated with a nominalised predicate (research question one formulated in the Introduction) as well as of the processes involved in exploiting this information for the linking of GenO and GenS arguments within a sentence (research question two). The following section 4.1 presents the SPR experiment tapping into reading time (RT) correlates of these two aspects; in section 4.1.7, I sketch a toy model which incorporates the main findings of this first experiment into a simple competition mechanism regulating the linking conflict. In section 4.2, I review the ERP experiment and map the most relevant results onto components of the toy model built on basis of the RT patterns. Note that the ERP data are analysed using a novel approach based upon the wavelet-based functional mixed model of Morris and Carroll (2006); section 4.2.6 provides the most important information about the present analysis approach; part II of this thesis discusses the relevant backgrounds as well as the details of the associated *wrapfmm* and *stepmom* software packages for the statistical environment R (R Development Core Team, 2013) in depth.

### 4.1 Self-paced reading experiment

As outlined in the preceding chapter, the SPR experiment presented below mainly aimed at assessing whether we can find RT correlates of the two aspects of interest during online sentence reading. The data obtained in the acceptability judgement study (see section 2.7) do not allow to draw any conclusions about the temporal locus of processes which may be

underlying the judgement patterns observed. In principle, the interaction witnessed may have been the result of some kind of delayed inferential processing occurring during or after sentence wrap-up; thus, the following SPR study tests for online analogues of the interaction of the 'AI' and 'Linking' variables at the genitive argument (and the immediately following region) to evaluate whether the effect arises immediately upon encountering the genitive. In addition, the online RT data allow to assess whether differences in lexical-semantic complexity of the nominalised predicate – as measured via its AI – correlate with RTs on the nominalisation and the subsequent region. The results of the SPR study will thus allow to draw a first picture of the temporal dynamics underlying the interplay of predicate-related lexical-semantic complexity and the resolution of predicate-induced argument linking conflicts within the current experimental paradigm; further, they will provide the behavioural basis for the subsequent ERP investigation. The predictions derived from the initial acceptability judgement study and the discussion of the psycholinguistic backgrounds in the preceding chapter are presented in section 4.1.3 below.

#### 4.1.1 Materials

Experimental items were constructed based upon the procedure applied in the acceptability rating study described in section 2.7.3.1, but were adapted to fit the requirements of the present SPR experiment. An item again consisted of short two-sentence stories with the first (context) sentence setting up the discourse context by describing an event or situation involving two participants. The second (continuation) sentence again represented a continuation of the story introduced by the context sentence, featuring the respective *-ung*-nominalisation, which was followed by a post-nominal genitive representing the subject of the context sentence (i.e. GenS) or its object (i.e. GenO). In contrast to the acceptability judgement study presented in the previous chapter, the NPs containing the nominalisation and the genitive argument were either embedded as sentence initial subjects or as post-verbal objects in the middle of the continuation sentence. In either case, the order of the verb and its argument NPs was always SVO and the post-verbal object was followed by different kinds of adjunct phrases (mostly PPs) comprising one to five words.

The story in (119) represents an example for an item (displaying the genitives occurring in the GenO and GenS condition) featuring a low-AI nominalisation (*Befragung – interrogation*), with the NP containing the nominalisation and the genitive argument embedded as the sentence-medial object of the continuation sentence; (120) represents an example of an item featuring a high-AI nominalisation (*Ermordung – assassination*) with the critical NP appearing as sentence-initial subject of the continuation sentence.

(119) Example story for *Befragung* (*interrogation*):

a. *Context sentence with 'role allocation':*

Der Staatsanwalt nahm den Angeklagten ins Kreuzverhör.  
 the.NOM prosecutor.NOM took the.ACC defendant.ACC into cross-examination  
 'The prosecutor cross-examined the defendant.'

b. *Continuation sentence:*

Der Verteidiger unterbrach die Befragung<sup>Nominalisation</sup> des<sup>GenArticle</sup>  
 the advocate interrupted the interrogation the.GEN  
 Angeklagt-en<sup>Genitive</sup> /Staatsanwalt-s<sup>Genitive</sup> mit<sup>Post1</sup> mehreren Einsprüchen.  
 defendant-GEN/prosecutor-GEN with repeated objections

'The advocate interrupted the interrogation of the defendant/prosecutor with repeated objections.'

(120) Example story for *Ermordung* (murder/assassination):

a. *Context sentence with 'role allocation':*

Der Attentäter tötete den Politiker während einer  
 the.NOM assassin.NOM killed the.ACC politician.ACC during an  
 Wahlveranstaltung.  
 election-event

'The assassin killed the politician during an election event.'

b. *Continuation sentence:*

Die Ermordung<sup>Nominalisation</sup> des<sup>GenArticle</sup> Politiker-s<sup>Genitive</sup> /Attentäter-s<sup>Genitive</sup>  
 the assassination the.GEN politician-GEN/assassin-GEN  
 verzögerte<sup>Post1</sup> die Wahl um einige Monate.  
 delayed the election for several months

'The assassination of the politician/assassin delayed the election for several months.'

Using 84 nominalisations whose base verbs were assigned an AI in the rating pretest described in section 2.7.2, experimental items were created in the following way: each of the 84 nominalisations was used for the construction of two *different* stories, one of which featured the critical nominalisation NP in sentence-initial subject position, the other one in sentence-medial object position, yielding a total of 168 stories, all of which had a 'GenO' and 'GenS' variant within the continuation sentences. Two experimental lists were created as follows: each list featured all of the 168 stories, with the 'Linking' variants of each story assigned to the lists in such a way that each participant was presented each nominalisation twice, once with a 'GenO' continuation and once with a 'GenS' continuation. The categorical factors 'Linking' ('GenO' and 'GenS' levels) and 'Embedding' ('Subject' and 'Object' levels) were thus balanced across lists and participants. Due to an encoding error, one item had to be removed prior to analysis, however. The two experimental lists were assigned to participants on an alternating basis and the order of presentation of stimuli within a list was randomised separately for each participant. Information on the nominalisations used is provided in table A.1 and the complete stimulus material is listed in Appendix C.

Note that the grammatical and lexical/conceptual criteria for the choice of the lexical fillers for the head nouns of the genitive NPs were the same as in the acceptability judgement study; further, the nominalisations' base verbs again did not occur directly in the context sentences

(as in the acceptability judgement experiment – see sections 2.7.2.1 and 2.7.3.1 for details). All verbs were presented in the past tense.

#### **4.1.2 Procedure**

The experiment was conducted using the Linger software by Doug Rohde. Reaction times were measured using Razer Arctosa keyboards with millisecond resolution.

At the beginning of each trial, the context and continuation sentences were masked completely by dashes; while spaces between words were masked in the context sentences, they were not masked in the continuation sentences. Upon the first press of the space bar, the complete context sentence was revealed. Upon the next press of the space bar, the context sentence was completely masked again and the first word of the continuation sentence was revealed. Each subsequent press revealed the following word and masked the previously visible word again, such that only one word at a time was visible in the continuation sentences. The sentences were presented in 18 point Courier font. Necessary line breaks in the continuation sentences never occurred within critical NPs. Participants were prompted to make a short break after every 20 items.

Participants were instructed to read the context sentences closely and make sure they understood their contents; there was no time limit for reading a context sentence. Further, they were told to read the continuation sentence word-by-word at a swift but nevertheless natural speed which allowed them to understand the sentences' contents.

At completion of the continuation sentence, participants were asked to rate the degree of acceptability of the second sentence as a coherent continuation of the context sentence. Instructions for the rating procedure were largely given as described in section 2.7.3.2 for the acceptability judgement study. In contrast to this previous study, however, ratings were not given as categorical 'Yes'/'No' responses, but provided on an integer scale ranging from 1–6, with '1' indicating complete unacceptability and a '6' straightforward acceptability of the second sentence as a coherent continuation. The switch from a binary to a graded scale was introduced in order to somewhat ease the rating task in the context of the primary reading time task. The ratings were provided by pressing the respective number keys. The participants were told to delay the decision about the ratings until after the completion of the continuation sentence and to focus on the sentences' contents during reading. They were also informed that there were no 'correct' or 'incorrect' ratings and were instructed to follow their own linguistic gut feeling in providing the judgements, to provide their responses quickly and to make use of the complete range of the six-point scale. In addition to the ratings, short 'Yes'/'No' comprehension questions were presented after a subset of 24 items to make sure participants were paying attention to the contents of the sentences. These targeted different aspects of the sentence pairs presented. The instructions contained a number of commented example items; before presentation of the experimental items, participants completed 4 practice trials. The experiment took place in a quiet room and each session lasted approximately 35–50 minutes.

### 4.1.3 Hypotheses

The analysis of the acceptability judgement data presented in chapter 2 revealed the following patterns (see sections 2.7.3.6 and 2.7.3.7): with GenOs, acceptability was high throughout the range of the AI variable, though a rise in AI still significantly increased acceptability of GenO linking. GenSs linked to nominalisations associated with a very low AI were equally acceptable as their GenO counterparts occurring with low-AI predicates; an increase in the AI was associated with a significant and sharp drop in acceptability, however, resulting in complete GenS unacceptability for nominalisations with a maximal AI. This pattern should reproduce with the acceptability rating data obtained in the SPR experiment, regardless of the adapted rating procedure (integer vs. binary scale).

As reviewed in section 3.2.1 above, a number of studies have found increased processing cost for higher lexical-semantic complexity of verbs (e.g. Gennari & Poeppel, 2003; McKoon & Love, 2011; McKoon & Macfarland, 2000, 2002; McKoon & Ratcliff, 2003) and some findings suggest that this cost arises immediately at encountering the verb during sentence parsing (e.g. Coll-Florit & Gennari, 2011; Gennari & Poeppel, 2003; McKoon & Love, 2011). In the discussion of these findings, I argued that they imply higher processing cost – i.e. longer RTs – with rising AI values in the current context, since verbs with less complex semantic structure (e.g. stative verbs) correspond to predicates situated at lower levels of the Affectedness Hierarchy of Beavers (2011) and those with more complex representations (e.g. eventive verbs) to higher levels. I also pointed out that these results may receive a complementary interpretation in terms of the number of lexical entailments associated with more complex verbs, a view which provides a direct conceptual link to Beavers' Affectedness Hierarchy and its connections to prototype-based theories of argument realisation (see Beavers, 2010; Dowty, 1991). This leads to hypothesis 1 about the effects of lexical semantic complexity on the predicate:

1. Syntactically transitive predicates implying a higher degree of affectedness for their objects map onto structurally more complex semantic representations; on a complementary view, their object arguments are associated with a higher number of Proto-Patient entailments. Thus, RTs for the eventive *-ung*-nominalisations and/or the following region ('Genitive Article') should increase with rising AI.

In the discussion of the acceptability patterns witnessed in chapter 2 (section 2.7.3.7), I argued that they can be explained in terms of semantic prominence of GenO and the closely related notions of GenO's prototypicality, which goes hand in hand with the semantic distance/proximity between the GenO and GenS co-arguments: when GenO is maximally prominent in terms of affectedness, it also corresponds best to the Proto-Patient role and the semantic distance to its GenS co-argument is maximal. As a consequence, more prominent/prototypical GenO arguments received slightly better ratings than less prominent/prototypical GenOs. In interpreting the acceptability patterns, I also speculated that they may be seen as the result of a competition of the two co-arguments for linking to the postnominal position, involving suppression of the competitor in dependence of semantic prominence of GenO.

On this view, as far as the patterns found for GenS linking are concerned, maximally prominent/prototypical GenOs can suppress their GenS co-arguments completely in the competition for linking to the postnominal argument position, while decreasing GenO prominence/prototypicality (and thus decreasing semantic distance to GenS) correlates with a loss of suppression potential for GenO and an increase of the ability to GenS to compete for linking to the postnominal slot. If such an explanation is on the right track and actually reflects processes taking place online as the genitive argument is encountered, rather than more general inferential processes occurring after the whole sentence has been read, RTs should follow the pattern stated in hypothesis 2:

2. If the acceptability patterns witnessed are the consequence of processes regulating the linking of GenS and GenO online, RTs for the genitive and/or the immediately following 'PostGenitive 1' regions should show a slight, but significant, *decrease* with rising AI in the 'GenO' condition; RTs in the 'GenS' condition, conversely, should exhibit the reverse effect, showing a clear and significant *increase* with rising AI. RTs at the low end of the AI scale should be about equal for the two genitive types, while they should differ maximally at the high end.

#### 4.1.4 Participants

20 native German speakers with a mean age of 21.7 (range 18–32, 5 male, 15 female) participated in the experiment in return for course credits or a monetary compensation of € 8,-, providing written and informed consent prior to participation.

#### 4.1.5 Data analysis

All analyses were performed using the statistical environment R (R Development Core Team, 2013). Details on the procedures for the RT and rating data are provided in the following.

#### Reading time data

The RT data were analysed for four regions, as indicated by the respective superscripts in examples (119) and (120), starting with the 'Nominalisation' region and the following 'GenArticle' (genitive article), 'Genitive' and 'Post1' (i.e. the first word after the critical NP with the nominalisation and genitive) regions. In four stories, the Post1 regions coincided with the final word of the sentence; to avoid confounds with sentence wrap-up effects, these were excluded from the analysis of the Post1 region.

The dependent variable used was the logarithm of the RTs of each single trial. Outliers were excluded for each region and participant separately following a three-step procedure which combined lenient a-priori trimming with trimming based on model criticism (a variant of the procedure discussed by Baayen & Milin, 2010): first, raw RTs below 150 ms or above 3000 ms were excluded, then any log-RTs more than 3 SDs above or below the

respective mean log-RT were excluded. Then, a first LMM was fitted to the log-RTs of these initial data sets using the `lmer` function of package *lme4* (Bates et al., 2014); based upon this initial model, data points with standardised residuals at a distance greater than 2.5 SDs from 0 were identified as outliers and removed using function `romr.fnc` from package *LMERConvenienceFunctions* (Tremblay & Ransijn, 2013). The data sets resulting from this procedure were subsequently used for the final model fitting. On average, this procedure resulted in the removal of 4.2% of trials from the full data sets.

All independent variables were centred on their mean, as indicated by the prefix 'c.'. In the 'Nominalisation' and 'GenArticle' regions preceding the presentation of the genitive, 'c.AI' was included as the sole fixed effect of interest; models for the following 'Genitive' and 'Post1' regions additionally included 'c.Linking' and the interaction of 'c.Linking' and 'c.AI' as predictors. Further, a number of control covariates were included in each region's model: 'c.Embedding' adjusted for embedding of the critical NP within sentence-initial subject or sentence-medial object position, and 'c.List.Pos' for the position of the respective trial within the experimental list; the respective word's length in characters was included as 'c.Word.Length' and the logarithm of its lemma frequency obtained from the *dlx*DB database (Heister et al., 2011) as 'c.log.Frequency'. In addition, to account for possible spillover effects from precedings words (Vasishth, 2006), the previous word's log-RT was added as predictor 'c.log.Spillover'. Since the word to be read in the 'GenArticle' region was always the same (*des*), this region's model did not include word length and frequency as control covariates.

Note that the `summary` method for LMMs fitted with `lmer` does not provide *p*-values for fixed effects (for discussion, see Baayen et al., 2008). To assess the significance of fixed effects, the *p*-values reported in the summaries below were computed using the `confint` and `glht` functions of package *multcomp* (Hothorn et al., 2008), which were also used to obtain confidence intervals (CIs) for the fixed effects. In addition, likelihood ratio tests were performed as a complementary way of assessing significance of fixed effects; these are reported and discussed only where they conflict with the *p*-values reported in the summary tables. Note that the summary tables below provide estimates on the log-scale; for better interpretability, all figures provided below are based on values back-transformed to the original millisecond scale.

Following the recommendations of Barr et al. (2013), maximal random effects structures for subjects and items were again used for model specification. Subject-related random effects comprised by-subject random intercepts and random slopes for the 'c.AI' predictor for all four regions; in addition, 'c.Linking' as well as its interaction with 'c.AI' were included as by-subject random slopes for the 'Genitive' and 'Post1' regions only. To model item-related variability, two grouping factors were used: to account for variability between the specific words read, 'Word' was used as a grouping factor with by-word random intercepts. In addition, 'Story' was again introduced as another grouping factor with by-story random intercepts and random slopes for the 'c.Linking' predictor, which varied within 'Story'. Since there was only one level of the 'Word' grouping factor for the 'GenArticle' region, no 'Word' random intercept was included for this region. Starting from these maximal models, the random effects structures were only omitted in the face of convergence problems or indications of

overparameterisation of the model (e.g. total correlation between random effects), starting with the removal of random correlation parameters, which were included by default in the initial maximal models (see Barr et al., 2013).

Planned post-hoc assessments comprised testing for simple slope effects of 'AI' for each level of 'Linking' for the 'Genitive' and 'Post1' regions by fitting separate LMMs for 'GenO' and 'GenS' subsets, taking the model specification of the final model for the complete data set as starting point and reducing the fixed and random effects parts as appropriate. Predictions for the continuous 'c.AI' predictor based upon the models' fixed effects estimates were computed along with 95% bootstrap-based prediction intervals as laid out in section 2.7.3.5 describing the analysis procedure for the initial acceptability judgement data. Model diagnostics included a check of the models' residuals as well as an assessment of outlying items, also as outlined in section 2.7.3.5.

## Rating data

The rating data were also analysed using standard (Gaussian) LMMs using the full data set, including 'c.AI' and 'c.Linkin' as well as their interaction as fixed effects; on the random effects side, by-subject random intercepts and by-subject random slopes for 'c.AI' and 'c.Linkin' as well as their interaction were included, plus by-story random intercepts and random slopes for 'c.Linkin'. All further steps were carried out as described for the RT data. Planned post-hoc assessments again comprised testing for simple slope effects of 'AI' at each level of 'Linkin'.

While it is widespread practice to analyse ordinal data with standard ANOVA or regression models, doing so is problematic for a number of reasons, including the violation of distributional assumptions about the data and the model's residuals (see Christensen, 2013a, for example). As a complementary check of the significance of fixed effects with more appropriate analysis methods, cumulative link mixed models ('CLMMs') were fitted to the rating data using the same fixed and random effects structure. The CLMMs were fitted using function `clmm` from the *ordinal* package (Christensen, 2013b); since assessment of significance of fixed effects converged well with the Gaussian LMMs in all cases, the results are not discussed further, but summary tables of the CLMMs are provided in Appendix C.

## 4.1.6 Results

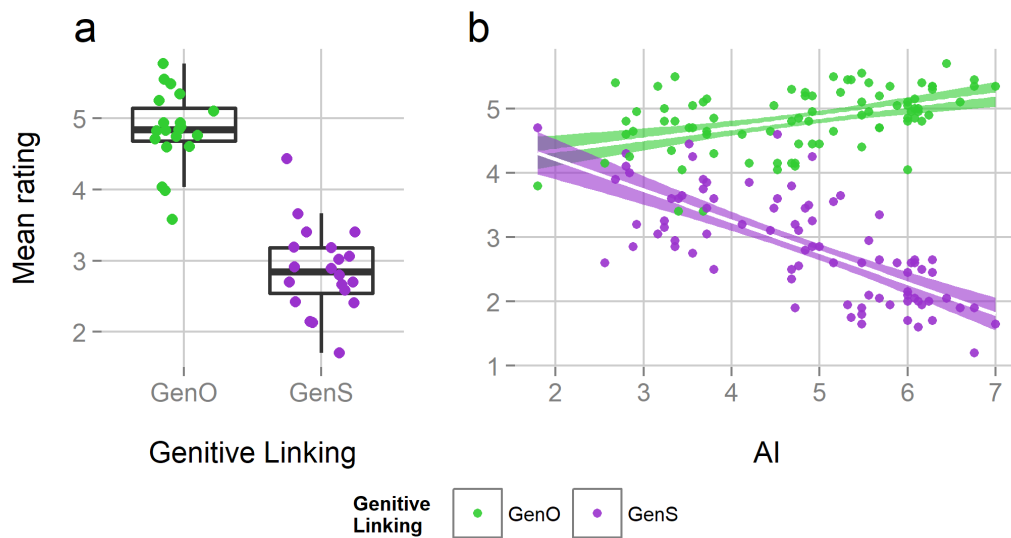
### Question response data

Question response accuracy was generally high, with a mean percentage of correct responses of 94.2% (range: 83.3% – 100%, SD = 5.1), suggesting that the participants paid close attention to the contents of the sentences.



## Rating data

The boxplots in figure 4.1a show the mean acceptability scores computed for each subject and 'Linking' level and illustrate that GenOs on average received higher acceptability ratings than GenSs (GenO: mean = 4.84, SD = 0.53; GenS: mean = 2.87, SD = 0.6). Figure 4.1b shows the mean acceptability rating per nominalisation at each level of 'Linking' over the range of 'AI' as a scatterplot with separate linear smooths for 'AI' for the 'GenO' and 'GenS' conditions. The smooth for the 'GenO' condition suggests good acceptability throughout the range of 'AI' and shows an increase with rising AI values. Conversely, in the 'GenS' condition, acceptability is as high as for 'GenO' continuations for nominalisations with a minimal AI and decreases as the AI increases.



**Figure 4.1:** Mean acceptability ratings for each 'Linking' level. a: Boxplot of mean acceptability ratings for each subject and 'Linking' level. b: Mean acceptability ratings for each nominalisation at the 'GenS' and 'GenO' levels of 'Linking' plus linear smooths for 'AI' at each level of 'Linking'.

The final model specification of the model for the complete data set in `lmer` notation is given in (121) and table 4.1 summarises the random and fixed effects estimates of the model.

```
(121) Rating ~ c.Linkin * c.AI +
      (1 + c.Linkin + c.AI | Subject) + (1 + c.Linkin | Story)
```

The model summarised in table 4.1 shows significant effects of 'c.Linkin' ( $B = -1.97$ ,  $SE = 0.21$ ,  $p < 0.001$ ), 'c.AI' ( $B = -0.15$ ,  $SE = 0.03$ ,  $p < 0.001$ ) as well as their interaction 'c.Linkin:c.AI' ( $B = -0.65$ ,  $SE = 0.06$ ,  $p < 0.001$ ).

Figure 4.2 summarises the predictions derived from model (121) together with their 95% bootstrap CIs. Figure 4.2a shows the predicted slopes of 'AI' at the 'GenO' and 'GenS'

#### Random effects

Groups	Name	Variance	SD	Corr.	
Story	(Intercept)	0.12	0.35		
	c.Linking	0.61	0.78	0.16	
Subject	(Intercept)	0.12	0.34		
	c.Linking	0.74	0.86	0.17	
	c.AI	0.01	0.08	-0.06	0.20
Residual		1.39	1.18		

#### Fixed effects

	Estimate	SE	Lower	Upper	t-value	Pr(> t )	
(Intercept)	3.86	0.08	3.69	4.02	46.15	<0.001	***
c.Linking	-1.97	0.21	-2.37	-1.57	-9.60	<0.001	***
c.AI	-0.15	0.03	-0.22	-0.09	-4.52	<0.001	***
c.Linking:c.AI	-0.65	0.06	-0.77	-0.54	-10.96	<0.001	***

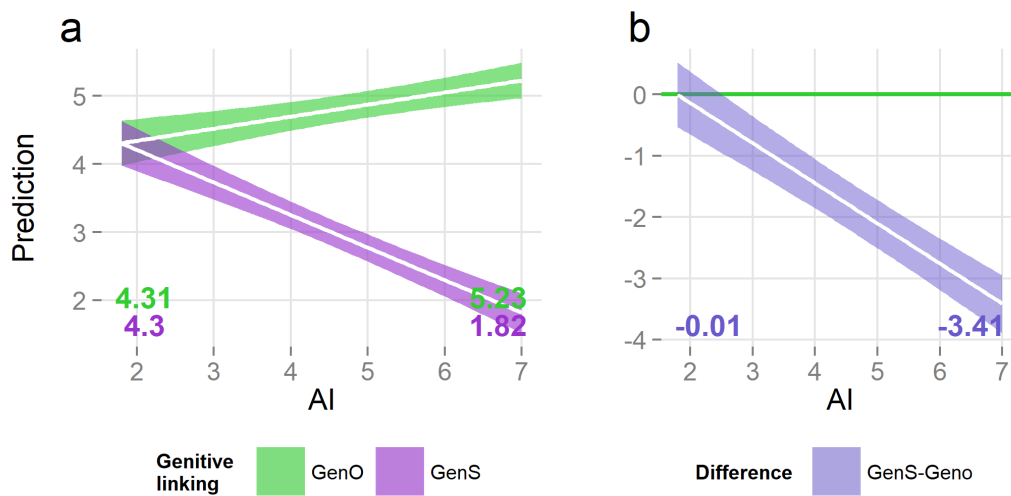
**Table 4.1:** Summary tables for the LMM for the acceptability rating data of the self-paced reading experiment.

levels of 'Linking'; it illustrates that the acceptability of a continuation sentence as a coherent continuation was equally high for GenO and GenS occurring with nominalisations associated with a minimal or low AI (estimate at the minimal AI value for GenO: 4.31; for GenS: 4.3). Acceptability decreased with increasing AI in the 'GenS' condition to minimal acceptability at the upper end of the 'AI' scale (estimate at the maximal AI value for GenS: 1.82), while it increased to maximal acceptability in the 'GenO' condition (estimate at the maximal AI value for GenO: 5.23). Figure 4.2b shows the estimates for the corresponding difference between 'GenS' and 'GenO' (i.e. GenS – GenO) across the 'AI' range with the corresponding bootstrap CIs.

Tables 4.2 and 4.3 summarise the post-hoc models fitted to test for simple slope effects of 'c.AI' specific to each level of 'Linking'; the respective model specifications are provided in (122).

- (122) a. *GenO-model*:  
Rating ~ c.AI +  
(1 + c.AI | Subject) + (1 | Story)
- b. *GenS-model*:  
Rating ~ c.AI +  
(1 + c.AI | Subject) + (1 | Story)

The model fitted to the 'GenO' subset (table 4.2) shows a significant effect of 'c.AI' ( $B = 0.18$ ,  $SE = 0.04$ ,  $p < 0.001$ ), i.e. a unit increase on the AI scale is associated with an increase of acceptability of 0.18 on the acceptability scale ranging from 1–6. Table 4.3 shows a significant effect of 'c.AI' in the model fitted to the 'GenS' subset ( $B = -0.48$ ,  $SE$



**Figure 4.2:** Predictions derived from model (121) with 95% bootstrap CIs. a: Predicted slopes of 'AI' at the 'GenO' and 'GenS' levels of 'Linking'. b: Predictions for GenS - GenO difference across the range of the AI variable. Numbers in the lower left and right corners indicate the estimates at the minimum and maximum values of the AI.

= 0.05,  $p < 0.001$ ); thus, in the 'GenS' condition, a unit increase on the AI scale is associated with a decrease of acceptability of  $-0.48$ .

## Reading time data

Figure 4.3 illustrates the RTs for a total of six consecutive regions, starting one region before the first critical region ('Nominalisation Article') and ending one region after the last critical one ('PostGenitive 2'), based upon aggregated log-RTs. The leftmost three panels of figure 4.3a show the mean overall RTs (based upon subject means) up to the 'Genitive Article' region, followed by the RTs for the 'GenO' and 'GenS' conditions from the 'Genitive' region onwards. The latter plots suggest that RTs were longer for 'GenS' continuations in the 'Genitive' region (log-RTs for 'GenO' vs. 'GenS' conditions: 6.03 vs. 6.1; on original ms scale: 416 ms vs. 444 ms) as well as in the 'PostGenitive 1' region (5.94 vs. 6.01; on original ms scale: 379 ms vs. 409 ms), with equal RTs in the following 'PostGenitive 2' region (5.95 vs. 5.95; on original ms scale: 382 ms vs. 384 ms). Figure 4.3b is based upon mean log-RTs for each nominalisation and shows linear smooths for 'AI' independent of the 'Linking' factor (first three panels) and separately for each of the 'GenS' and 'GenO' levels (last three panels). While the smooths in the 'Nominalisation' and 'Genitive Article' regions do not reveal any obvious tendencies, the patterns in the 'Genitive' and 'PostGenitive 1' regions are consistent with the predicted interaction, showing longer RTs with increasing 'AI' in the 'GenS' condition and the reverse pattern in the 'GenO' condition, with roughly equal RTs at

Random effects

Groups	Name	Variance	SD	Corr.
Story	(Intercept)	0.25	0.50	
Subject	(Intercept)	0.25	0.50	
	c.AI	0.00	0.05	-0.38
Residual		1.22	1.10	

Fixed effects

	Estimate	SE	Lower	Upper	t-value	Pr(> t )	
(Intercept)	4.84	0.12	4.6	5.08	39.73	<0.001	***
c.AI	0.18	0.04	0.1	0.26	4.41	<0.001	***

**Table 4.2:** LMM results for the GenO subset of the self-paced reading acceptability data.

Random effects

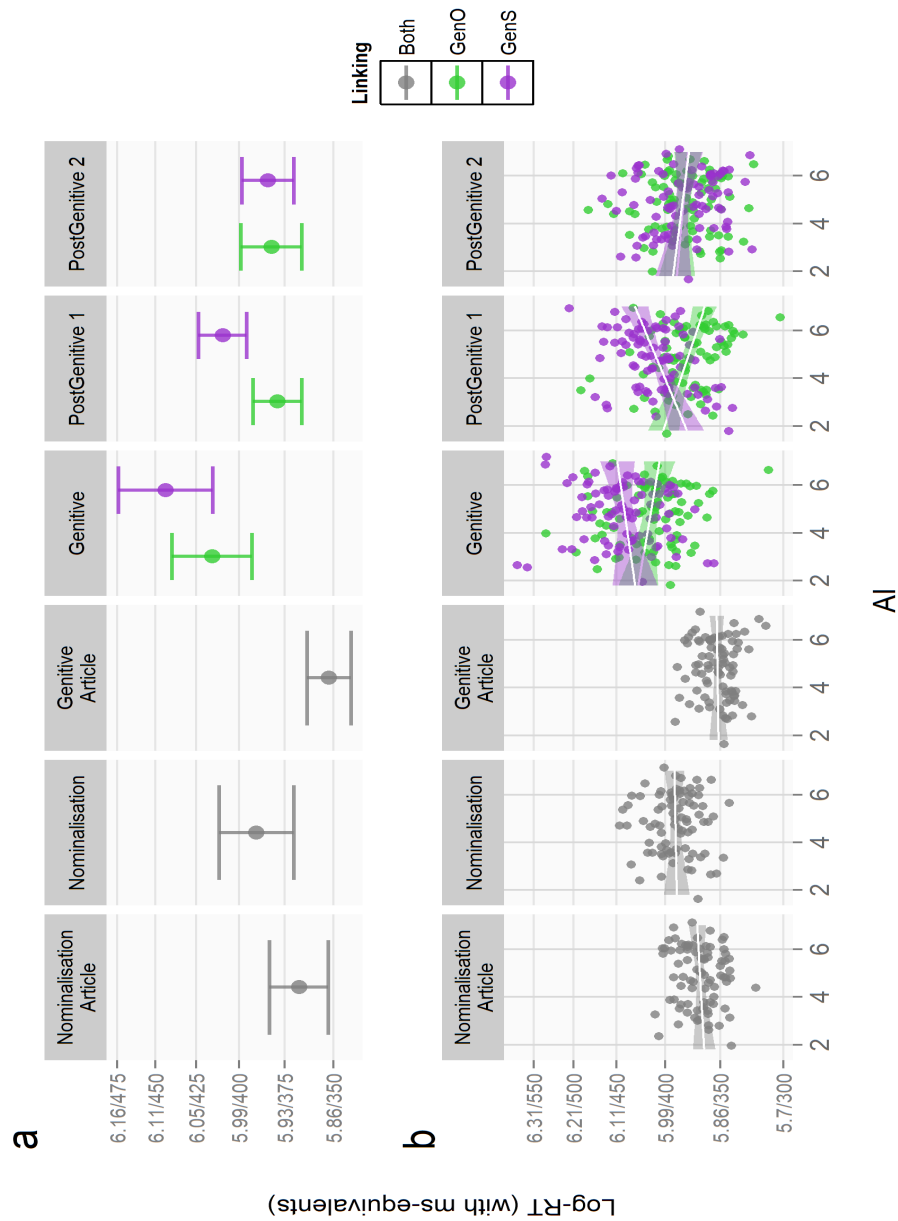
Groups	Name	Variance	SD	Corr.
Story	(Intercept)	0.30	0.55	
Subject	(Intercept)	0.35	0.59	
	c.AI	0.02	0.13	-0.14
Residual		1.55	1.25	

Fixed effects

	Estimate	SE	Lower	Upper	t-value	Pr(> t )	
(Intercept)	2.87	0.14	2.59	3.15	20.29	<0.001	***
c.AI	-0.48	0.05	-0.58	-0.38	-9.30	<0.001	***

**Table 4.3:** LMM results for the GenS subset of the self-paced reading acceptability data.

the low end of the 'AI' scale. This pattern is absent in the 'PostGenitive 2' region.



**Figure 4.3:** Reading times for six regions. a: Mean overall-RTs (first three panels) and RTs for each 'Linking' condition (last three panels). Error bars represent simple standard error of the mean based upon subject means. b: Mean RTs for each nominalisation across both levels of 'Linking' (first three panels) and for the 'GenS' and 'GenO' levels (last three panels) arranged by AI, plus linear smooths. All quantities were computed from log-RTs; for better interpretability, the labels on the y-axes provide the log-RTs as well as the corresponding values on the original ms scale.

The model specifications of the models fitted to the RT data of the 'Nominalisation' and

'Genitive Article' regions are given in (123). Tables 4.4 and 4.5 summarise the respective models.

- (123) a. *'Nominalisation' region:*  
 $\log.RT \sim c.AI + c.Embedding + c.Word.Length + c.log.Frequency + c.List.Pos + c.log.Spillover + (1 | Subject) + (0 + c.AI | Subject) + (1 | Story) + (1 | Word)$
- b. *'Genitive Article' region:*  
 $\log.RT \sim c.AI + c.Embedding + c.List.Pos + c.log.Spillover + (1 + c.AI | Subject) + (1 | Story)$

Random effects

Groups	Name	Variance	SD
Story	(Intercept)	0.000	0.000
Word	(Intercept)	0.000	0.008
Subject	c.AI	0.000	0.000
Subject.1	(Intercept)	0.047	0.217
Residual		0.052	0.227

Fixed effects

	Estimate	SE	Lower	Upper	t-value	Pr(> t )	
(Intercept)	5.967	0.049	5.871	6.062	122.52	<0.001	***
c.AI	0.001	0.003	-0.005	0.008	0.43	0.668	
c.Embedding	-0.016	0.008	-0.032	0.000	-1.98	0.048	*
c.Word.Length	0.018	0.002	0.013	0.023	7.61	<0.001	***
c.log.Frequency	-0.005	0.003	-0.010	0.001	-1.62	0.105	
c.List.Pos	-0.002	0.000	-0.002	-0.002	-20.60	<0.001	***
c.log.Spillover	0.165	0.013	0.140	0.190	12.83	<0.001	***

**Table 4.4:** LMM results for the RT-data of the 'Nominalisation' region.

While the effects of the control predictors are generally consistent with the respective expectations (e.g. longer RTs with increasing word length and spillover; shorter RTs with increasing word frequency and position within the experimental list), the 'c.AI' predictor fails to reach significance in the 'Nominalisation' region ( $B = 0.0015$ ,  $SE = 0.0034$ ,  $p = 0.66799$ ) as well as in the 'Genitive Article' region ( $B = -0.0017$ ,  $SE = 0.0036$ ,  $p = 0.63858$ ).

The models fitted to the data of the 'Genitive' and 'PostGenitive 1' regions are given in (124); tables 4.6 and 4.7 summarise the fixed and random effects estimates.

- (124) a. *'Genitive' region:*

Random effects

Groups	Name	Variance	SD	Corr.
Story	(Intercept)	0.001	0.029	
Subject	(Intercept)	0.017	0.129	
	c.AI	0.000	0.007	0.05
Residual		0.034	0.184	

Fixed effects

	Estimate	SE	Lower	Upper	t-value	Pr(> t )	
(Intercept)	5.864	0.029	5.807	5.921	201.32	<0.001	***
c.AI	-0.002	0.004	-0.009	0.005	-0.47	0.639	
c.Embedding	-0.055	0.008	-0.071	-0.040	-6.99	<0.001	***
c.List.Pos	-0.001	0.000	-0.001	-0.001	-19.22	<0.001	***
c.log.Spillover	0.081	0.010	0.062	0.099	8.45	<0.001	***

**Table 4.5:** LMM results for the RT-data of the 'Genitive Article' region.

```
log.RT ~ c.Linking * c.AI + c.Embedding + c.List.Pos +
c.Word.Length + c.log.Frequency + c.log.Spillover +
(1 | Subject) + (0 + c.Linking | Subject) +
(0 + c.AI | Subject) + (0 + c.Linking:c.AI | Subject) +
(1 | Story) + (0 + c.Linking | Story) + (1 | Word)
```

b. 'PostGenitive 1' region:

```
log.RT ~ c.Linking * c.AI + c.Embedding + c.List.Pos +
c.Word.Length + c.log.Frequency + c.log.Spillover +
(1 + c.Linking * c.AI | Subject) + (1 | Story) +
(0 + c.Linking | Story) + (1 | Word)
```

Summary table 4.6 for the 'Genitive' region discloses significant effects of the 'c.Linking' predictor ( $B = 0.058$ ,  $SE = 0.018$ ,  $p = 0.0012$ ) and the 'c.Linking:c.AI' interaction ( $B = 0.021$ ,  $SE = 0.011$ ,  $p = 0.048$ ); the latter, however, fails to reach significance in the complementary likelihood ratio test performed ( $p_{\chi^2} = 0.0515$ ). The estimates for the subsequent 'PostGenitive 1' region likewise include significant effects of the 'c.Linking' predictor ( $B = 0.071$ ,  $SE = 0.019$ ,  $p < 0.001$ ) and a pronounced 'c.Linking:c.AI' interaction ( $B = 0.045$ ,  $SE = 0.009$ ,  $p < 0.001$ ).

Figure 4.4 shows the predictions for the RTs derived from model (124b) together with their 95% bootstrap CIs for the 'PostGenitive 1' region, in which the predicted interaction effect of 'c.Linking' and 'c.AI' emerged most clearly. Figure 4.4a illustrates that 'c.AI' had a facilitatory effect in the 'GenO' condition, as the RT decreased from 407 ms at the very low end of the AI scale to 362 ms at its maximum. Conversely, in the 'GenS' condition, RT increased as a function of the Affectedness Index, from 382 ms at minimal AI values to 430 ms at the very high end of the AI scale. Figure 4.4b suggests that RTs in the 'GenS' condition were actually 25 ms faster with nominalisations associated with a minimal AI than in the 'GenO' condition,

Random effects

Groups	Name	Variance	SD
Story	c.Linking	0.002	0.041
Story.1	(Intercept)	0.002	0.040
Word	(Intercept)	0.000	0.021
Subject	c.Linking:c.AI	0.001	0.030
Subject.1	c.AI	0.000	0.012
Subject.2	c.Linking	0.004	0.066
Subject.3	(Intercept)	0.070	0.264
Residual		0.067	0.260

Fixed effects

	Estimate	SE	Lower	Upper	t-value	Pr(> t )	
(Intercept)	6.062	0.059	5.946	6.179	102.02	<0.001	***
c.Linking	0.058	0.018	0.023	0.094	3.24	0.001	**
c.AI	-0.001	0.005	-0.011	0.010	-0.11	0.910	
c.Embedding	-0.094	0.011	-0.117	-0.072	-8.31	<0.001	***
c.List.Pos	-0.002	0.000	-0.002	-0.002	-21.14	<0.001	***
c.Word.Length	0.021	0.003	0.016	0.027	7.40	<0.001	***
c.log.Frequency	-0.008	0.003	-0.015	-0.001	-2.33	0.020	*
c.log.Spillover	0.119	0.015	0.090	0.148	7.98	<0.001	***
c.Linking:c.AI	0.021	0.011	0.000	0.042	1.98	0.048	*

**Table 4.6:** LMM results for the RT-data of the 'Genitive' region.

whereas 'GenO' continuations show a processing advantage of 68 ms at the upper end of the AI scale.

The models fitted to the 'GenO' and 'GenS' subsets to test for the simple slopes of 'c.AI' in the 'PostGenitive 1' region are stated in (125), with the respective summaries provided in tables 4.8 and 4.9. The LMMs show significant effects of 'c.AI' for the 'GenO' subset ( $B = -0.023$ ,  $SE = 0.005$ ,  $p < 0.001$ ) as well as for the 'GenS' subset ( $B = 0.023$ ,  $SE = 0.007$ ,  $p = 0.001004$ ).

- (125) a. *Model for 'GenO' subset of the 'PostGenitive 1' region:*  
 $\log.RT \sim c.AI + c.Embedding + c.List.Pos +$   
 $c.Word.Length + c.log.Frequency + c.log.Spillover +$   
 $(1 | Subject) + (1 | Story) + (1 | Word)$
- b. *'GenS' subset:*  
 $\log.RT \sim c.AI + c.Embedding + c.List.Pos +$   
 $c.Word.Length + c.log.Frequency + c.log.Spillover +$   
 $(1 + c.AI | Subject) + (1 | Story) + (1 | Word)$



#### Random effects

Groups	Name	Variance	SD	Corr.
Story Story.1 Word Subject	c.Linking	0.002	0.043	
	(Intercept)	0.000	0.000	
	(Intercept)	0.001	0.026	
	(Intercept)	0.020	0.140	
	c.Linking	0.005	0.073	-0.04
	c.AI	0.000	0.008	0.34 0.84
	c.Linking:c.AI	0.001	0.022	0.59 0.73 0.96
Residual		0.060	0.245	

#### Fixed effects

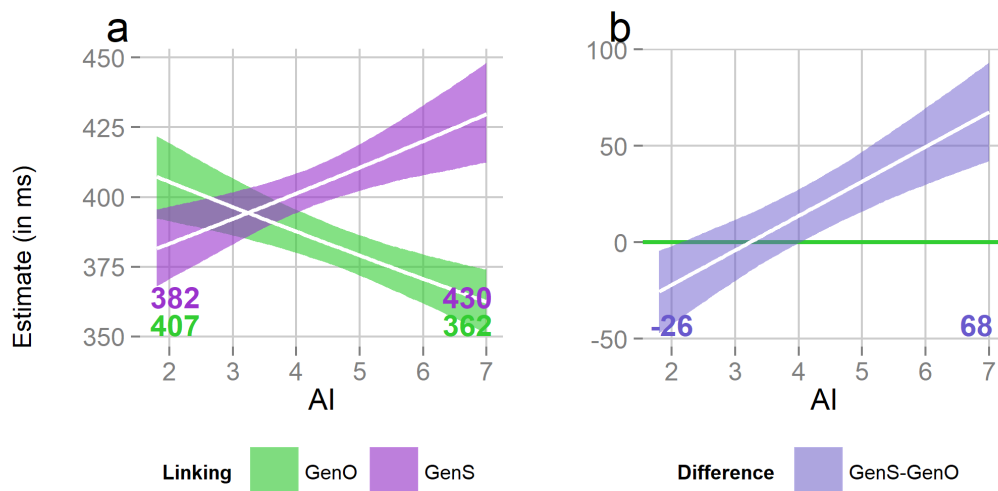
	Estimate	SE	Lower	Upper	t-value	Pr(> t )	
(Intercept)	5.977	0.032	5.915	6.040	187.62	<0.001	***
c.Linking	0.071	0.019	0.034	0.108	3.74	<0.001	***
c.AI	0.000	0.004	-0.008	0.008	0.00	1.000	
c.Embedding	-0.053	0.022	-0.095	-0.011	-2.47	0.014	*
c.List.Pos	-0.002	0.000	-0.002	-0.002	-20.11	<0.001	***
c.Word.Length	0.012	0.004	0.005	0.020	3.32	<0.001	***
c.log.Frequency	-0.007	0.005	-0.017	0.003	-1.37	0.170	
c.log.Spillover	0.061	0.011	0.039	0.083	5.45	<0.001	***
c.Linking:c.AI	0.045	0.009	0.027	0.063	4.95	<0.001	***

**Table 4.7:** LMM results for the RT-data of the 'PostGenitive 1' region.

## 4.1.7 Discussion

### Rating data

The results for the rating data are fully consistent with the findings of the acceptability rating study presented in the previous chapter and replicate the patterns identified. This independent replication suggests that the acceptability results generalise beyond the particular methodology (offline vs. online sentence processing) and rating scale used (binary judgements vs. graded ratings). The interaction of 'Linking' and 'AI' shows the predicted form and the pattern is near-identical to that witnessed in the rating study, as a comparison of figures 4.2 and 2.3 confirms: equally good acceptance of GenO and GenS with minimum-AI nominalisations, a slight and significant increase of acceptability for GenOs with increasing AI and a sharp decline in acceptability with rising AI for GenS continuations, resulting in near-bottom acceptability for GenSs linked to maximum-AI predicates. The current results are thus also compatible with the interpretation of the acceptability patterns I suggested previously and the role of semantic prominence/prototypicality of the GenO argument, which not only accounts for the slight increase in acceptability of GenO with rising AI values, but also for the sharp acceptability drop for GenSs linked to higher AI-nominalisations via the potential of GenO to suppress its GenS co-argument in the competition for linking to the postnominal



**Figure 4.4:** Predictions derived for the 'PostGenitive 1' region from model (124b) with 95% bootstrap CIs. a: Predicted slopes of 'AI' at the 'GenO' and 'GenS' levels of 'Linking'. b: Predictions for GenS - GenO difference across the range of AI. Numbers in the lower left and right corners indicate the estimates at the minimum and maximum values of AI.

genitive slot.

## RTs on the 'Nominalisation' and 'Genitive Article' regions

The analysis of the RTs for the 'Nominalisation' and the following 'Genitive Article' regions failed to find any effects of 'AI'. This lack of an effect contradicts the prediction of longer RTs with increasing AI (hypothesis 1 above), which was made on basis of the findings of a number of studies which have found evidence for processing cost associated with lexical-semantic complexity of verbs (e.g. Gennari & Poeppel, 2003; McKoon & Love, 2011; McKoon & Macfarland, 2000, 2002; McKoon & Ratcliff, 2003).

One possible explanation for the failure to find an associated effect may be that the nominalisations – or their base verbs – were primed by the paraphrases or synonyms which were presented in the initial context sentence. Such priming of the nominalisations or base verbs and the associated semantic structure may have eliminated any effects related to the AI. In fact, such a possible priming of the target nominalisations is an inherent weakness of the current experimental design. Unfortunately, it is also an inevitable flaw of the study, since the target linking configuration has to be fixed prior to presenting the nominalised predicate and the genitive argument in the continuation sentence.

Another plausible factor may have imageability – Gennari and Poeppel (2003), for example, noted that the eventive verbs applied in their study were rated more imageable than the stative verbs used, which creates a situation in which imageability potentially counteracts the in-

Random effects

Groups	Name	Variance	SD
Story	(Intercept)	0.000	0.000
Word	(Intercept)	0.002	0.040
Subject	(Intercept)	0.021	0.144
Residual		0.051	0.226

Fixed effects

	Estimate	SE	Lower	Upper	t-value	Pr(> t )	
(Intercept)	5.943	0.033	5.878	6.009	178.27	<0.001	***
c.AI	-0.023	0.005	-0.033	-0.013	-4.55	<0.001	***
c.Embedding	-0.051	0.030	-0.109	0.007	-1.74	0.083	
c.List.Pos	-0.002	0.000	-0.002	-0.001	-14.33	<0.001	***
c.Word.Length	0.013	0.005	0.003	0.023	2.47	0.013	*
c.log.Frequency	-0.001	0.007	-0.015	0.012	-0.17	0.864	
c.log.Spillover	0.091	0.015	0.062	0.121	6.03	<0.001	***

**Table 4.8:** LMM results for the post-hoc RT analysis of the 'GenO' subset of the 'PostGenitive 1' region.

creased processing cost caused by lexical-semantic complexity, since more imageable words usually have a processing advantage over less imageable ones. In the study of Gennari and Poeppel, eventive verbs still took longer to process than stative verbs, but imageability may have contributed to shrinking a possible effect of the AI in the current experiment. A similar issue arises with the temporal duration of the events denoted by the nominalisations' base verbs, since Coll-Florit and Gennari (2011) showed that durativity incurs processing cost in the processing of verbs and higher duration ratings correlated positively with reading times. While durativity probably varied throughout the range of the 'AI' covariate, a brief look at the five base verbs with the lowest (*admire/tail/interview, interrogate/greet/advise*) and highest (*decapitate/kill/murder/remove/cure*) AI values suggests that it may nevertheless have been an additional counteracting factor, since these lower-AI items mostly imply longer temporal durations than the high-AI verbs, at least in my intuition. If this brief and informal check is representative for the set of items used, then higher durativity associated with lower-AI verbs could have additionally cancelled out effects induced by affectedness. Thus, these two factors – imageability and durativity – may have jointly contributed to neutralising possible effects of affectedness.

Since the LMMs defined did not test for non-linear effects of continuous predictors, another possibility may be that the linear terms used for modelling effects of 'c.AI' did not adequately capture potentially non-linear patterns in the data. Figure 4.5 shows the mean log-RTs for each nominalisation in the 'Nominalisation' and 'Genitive Article' again, as already displayed in figure 4.3 above, with non-linear smooths based upon generalised additive models (GAMs). An inspection of these smooths does not suggest the presence of a pronounced non-linearity, however.

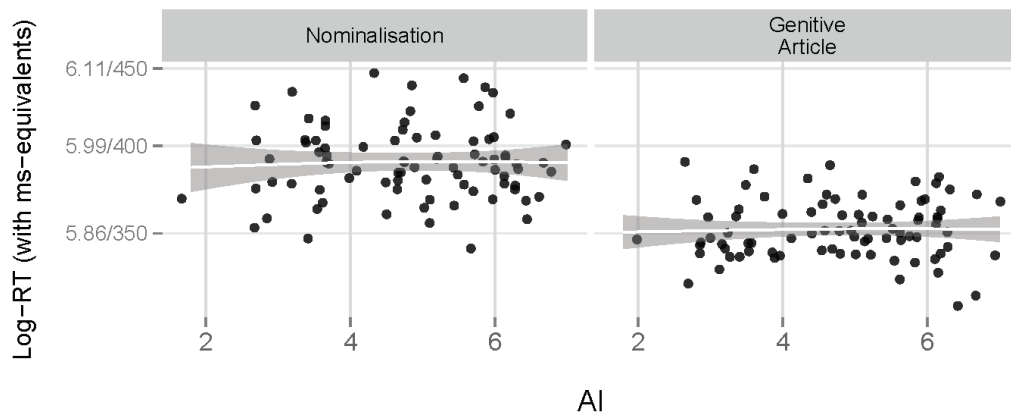
Random effects

Groups	Name	Variance	SD	Corr.
Story	(Intercept)	0.001	0.028	
Word	(Intercept)	0.000	0.000	
Subject	(Intercept)	0.021	0.147	
	c.AI	0.000	0.018	0.75
Residual		0.068	0.261	

Fixed effects

	Estimate	SE	Lower	Upper	t-value	Pr(> t )	
(Intercept)	6.013	0.034	5.947	6.078	179.36	<0.001	***
c.AI	0.023	0.007	0.009	0.037	3.29	0.001	**
c.Embedding	-0.057	0.027	-0.111	-0.004	-2.10	0.036	*
c.List.Pos	-0.002	0.000	-0.002	-0.002	-14.25	<0.001	***
c.Word.Length	0.014	0.005	0.005	0.023	2.92	0.004	**
c.log.Frequency	-0.010	0.006	-0.022	0.003	-1.55	0.122	
c.log.Spillover	0.032	0.017	0.000	0.064	1.94	0.053	

**Table 4.9:** LMM results for the post-hoc RT analysis of the 'GenS' subset of the 'PostGenitive 1' region.



**Figure 4.5:** Mean RTs for each nominalisation arranged by AI plus non-linear smooths based on GAMs.

Finally, the absence of an effect of lexical-semantic complexity may be interpreted as being in line with those previous studies which failed to find evidence for processing cost induced by structurally complex verbal representations (e.g. Fodor et al., 1980; Rayner & Duffy, 1986). In particular, Manouilidou and de Almeida (2013) failed to find evidence for increased processing cost for change of state verbs in an SPR experiment, while factors related to aspects of the argument structure of verbs actually affected reading times; these authors suggest that their findings may be understood as doubting the prominent role ascribed to lexical-

semantic complexity in verb processing. However, since it is not clear to me whether or in how far properties such as word length and potentially counteracting effects of imageability and durativity were accounted for in this study, it is hard to pinpoint the source of these null results.

## RTs on the 'Genitive' and 'PostGenitive 1' regions

The RTs for the 'Genitive' and 'PostGenitive 1' regions largely confirmed the predictions of hypothesis 2; the LMM analysis revealed a significant interaction of the 'c.AI' and 'c.Linking' predictors which exhibited the predicted form: roughly equal RTs at the low end of the AI-scale, and while an increase in the 'c.AI' predictor correlated with *shorter* RTs in the 'GenO' condition, it correlated with *longer* RTs in the 'GenS' condition. This effect emerged most clearly in the 'PostGenitive 1' region, i.e. the immediate spillover region of the 'Genitive' region; the planned post-hoc assessments of the simple slopes of the 'c.AI' predictor revealed a significant negative slope for GenO continuations and a significant positive slope for their GenS counterparts.

The locus of this effect is consistent with interpreting it as a reflection of processes regulating the integration of GenO and GenS online, as the relevant genitive is encountered. On this view, the acceptability judgements are at least partially determined by the outcome of these online processes. If the interpretation of the acceptability ratings offered previously in section 2.7.3.6 is on the right track, the present patterns may be seen as processing correlates of diagnosing and/or resolving prominence-induced linking conflicts. Following my previous arguments, the underlying processes may stem from the attempt to immediately link the genitive argument encountered either as GenO or GenS to establish a predicate-argument complex which is in line with the linking established in the context sentence, which acts as the relevant target linking configuration. As the rating patterns of the present and the previous experiment suggest, the decisive factor is the AI associated with the base verb of the respective nominalisation, which defines the semantic prominence of the GenO argument.

The present experiment has been successful in demonstrating that the processes which regulate the prominence-induced linking conflicts are deployed immediately as the relevant genitive argument is integrated. This finding provides the basis for the attempt to isolate electrophysiological correlates in the following ERP experiment. However, so far I have been somewhat vague about the functional interpretation of the underlying processes; while hypothesis 2 about the interaction effect indexing linking-related processes was largely motivated by reference to the prominence-induced 'suppression potential' of GenO and GenS – based on previously obtained rating patterns –, I have not addressed the details of potential processing steps involved and their functional interpretations so far. In the following, I will sketch a 'toy model' of the competitive linking process based upon the acceptability and current RT results.

In general, one of the main tasks of the underlying processes would be to 'mediate' between the target linking (established within the discourse representation) and the encountered argument online, via the semantic prominence status of GenO. One assumption I will make in

building the toy model is that GenO and GenS arguments of *-ung*-nominalisations are allocated an *inherent suppression potential* (ISP) as follows: first, the total amount of available ISP is limited and fixed; the GenO and GenS arguments of a syntactically transitive *-ung*-nominalisation compete for ISP from this pool. Second, the allocation of ISP to any GenO/GenS argument is regulated by GenO's semantic prominence as defined by the AI: the higher GenO's AI/semantic prominence, the higher its ISP and the lower its GenS co-argument's ISP; with decreasing AI/semantic prominence of GenO, on the other hand, the proportion of ISP allocated to GenO shrinks, while the amount of ISP allocated to GenS increases. Thus, the amount of ISP of a GenS is the proportion of ISP left over in the ISP pool after GenO has been allocated its share of ISP. For the sake of concreteness and illustration, I will assume that the general ISP pool contains a fixed amount of 1 and the quantity of ISP attributed to any genitive argument is measured on a continuous scale ranging from 0 to 1; a GenO of a maximum-AI nominalisation is allocated the full amount of 1 ISP, leaving 0 ISP for its GenS co-argument; further, I will assume that a GenO associated with a minimum-AI predicate receives 0.5 ISP from the pool, leaving the same amount of 0.5 ISP for GenS. The ISP may be conceptually thought of as a lexically determined processing resource which reflects inherent properties of the respective predicate's lexical-semantic structure.

The ISP is one component of the simple competitive linking mechanism proposed in the following. As discussed below, when either a GenO or GenS argument of an *-ung*-nominalisation is encountered in a sentence, the suppression resources of either argument are pitted against each other; this implies that an attempt to integrate a given GenO/GenS always involves mobilising the suppression resources of both arguments at the same time, even if one of them is not realised within the sentence and thus only present implicitly. Importantly, since the total amount of ISP available for allocation is fixed and always sums to 1 across GenO and GenS, it does not incur varying (or additional) processing costs across different constructions (i.e. experimental items in the current context).

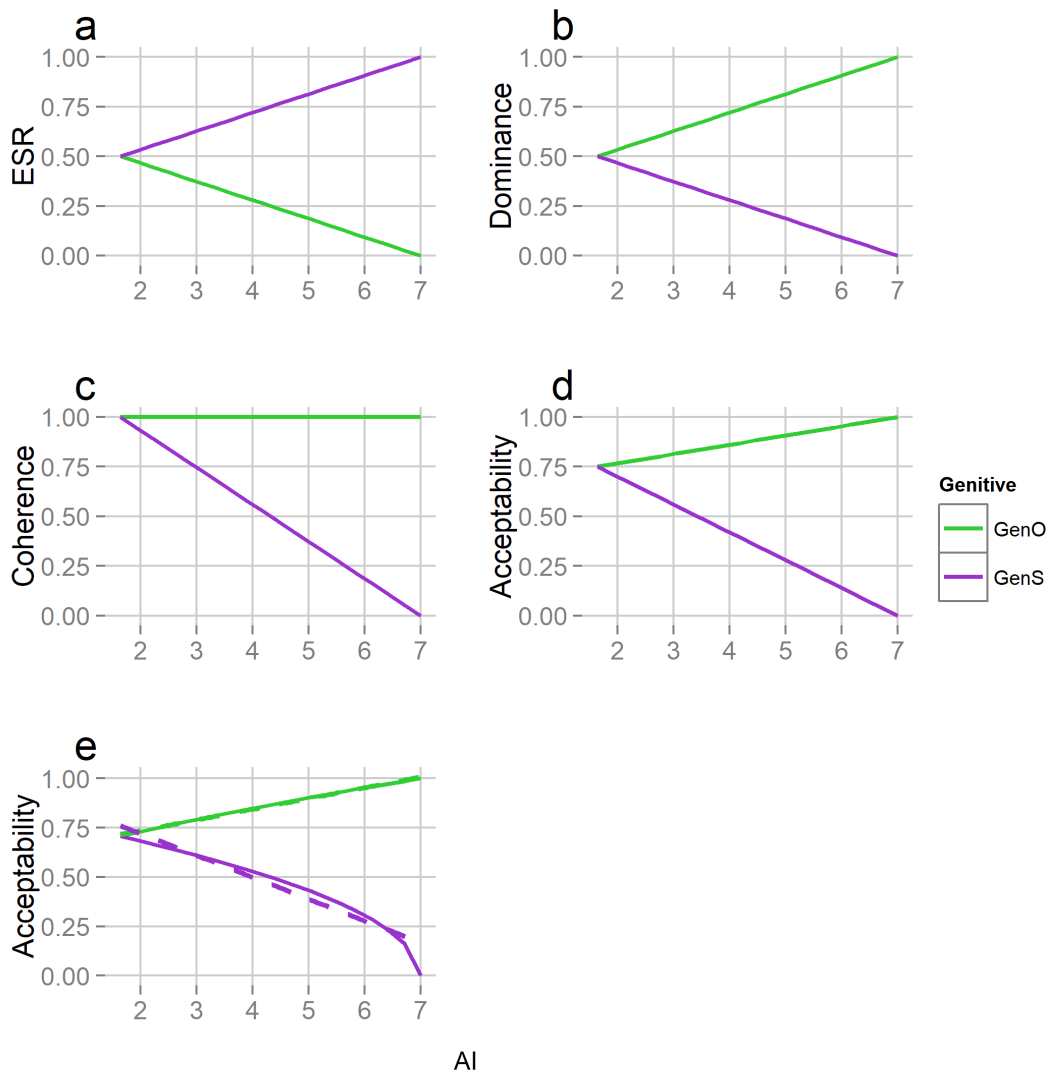
An additional building block of the proposed mechanism is based upon the assumption that the parser does its best to attempt and successfully link the encountered genitive argument to the nominalised predicate in order to maintain coherence with the target linking established in the context sentence. Within the toy model proposed, 'doing its best' amounts to make *additional* suppression resources available for the specific GenO or GenS argument encountered. Specifically, I propose that the parser fills up the encountered genitive NP's suppression potential to the maximum possible; since an ISP of 1 reflects a maximally prominent GenO argument, I assume that the parser mobilises enough additional suppression resources to boost the currently processed genitive argument's suppression potential to the maximum of 1. In the following, I will refer to these additional processing resources as *external suppression resources* (ESR).

The amount of ESR mobilised for any given encountered genitive argument is thus calculated simply as  $1 - ISP$ . Figure 4.6a illustrates the amount of ESR allocated to a genitive NP encountered as a function of the AI and the NP's syntactic status (GenO vs. GenS): for GenOs, it is minimal with a maximal AI and increases up to 0.5 for GenOs with a minimal AI; for a GenS of a minimum-AI nominalisation, the ESR amounts to 0.5, while it increases

up to 1 for a GenS associated with a maximum-AI predicate. Crucially, in contrast to the ISP, the mobilisation of ESR thus incurs additional processing cost which varies across items; the additional resources invested should thus affect RTs in the 'Genitive' or 'PostGenitive 1' regions in a way consistent with the pattern in figure 4.6a. A brief comparison with the model predictions of the RT pattern for the 'PostGenitive 1' region in figure 4.4a suggests that the hypothesised impact of the ESR indeed approximates the data reasonably well: the RTs are approximately equal at the low end of the AI-scale, the directions of the effects match the predictions and the slopes for the GenO and GenS conditions are roughly symmetric, in accordance with the ESR component of the toy model.

The next step within the toy model consists of determining the extent to which the argument encountered dominates its implicit co-argument, which is achieved by simply subtracting the co-argument's *total suppression potential* (TSP) from its own; given that the encountered argument's TSP always equals 1, and its co-argument's TSP corresponds to the initial ISP allocated, the dominance of the genitive to be integrated over its co-argument is given by  $1 - ISP_{coarg}$ . Figure 4.6b illustrates the dominance of a GenO or GenS in dependence of the AI. Assuming that the resulting dominance reflects the degree to which the co-argument can be suppressed in the competition for linking to the postnominal genitive slot, the dominance pattern should correlate with the acceptability ratings: a GenS associated with a maximum-AI nominalisation results in a dominance of 0, suppression of its GenO co-argument thus fails completely and inacceptability ensues; a GenO located at the upper end of the AI scale, on contrast, is associated with a dominance of 1, it can thus completely suppress its GenS co-argument, yielding optimal acceptability; at the lower end of the AI scale, GenO and GenS dominate their respective co-arguments by the same amount (0.5), predicting equal acceptability with low-AI predicates.

A look back at the rating patterns illustrated by figures 4.2a and 2.3a suggests that these specific predictions are indeed born out. One conspicuous difference between the patterns expected on basis of the dominance and the empirical acceptability patterns, however, lies in the slopes, which should be roughly symmetric according to the toy model, but are actually asymmetric, suggesting that the mapping from dominance to the eventual rating is not one-to-one. The empirical patterns may be derived from dominance in a number of ways; one possibility may be that genitive NPs associated with a dominance value at or above a certain threshold are actually perceived as fully coherent with the target linking established in the first context sentence, while the coherence of genitives associated with a dominance index below this threshold declines linearly as a function of the AI. Setting the threshold at 0.5, the resulting coherence perception is illustrated by figure 4.6c. The final acceptability ratings may then be modelled as the mean of the dominance and coherence values; figure 4.6d suggests that the acceptability ratings predicted by this approach closely correspond to the empirical findings. Alternatively, the mapping from dominance to acceptability may follow a non-linear function: 4.6e shows the slopes resulting from taking the square root of the dominance indices (solid lines) as well as linear smooths of these square-rooted values (broken lines). The linear smooths again closely approximate the empirical acceptability patterns. This mapping from dominance to a rating of an item's acceptability may in principle take



**Figure 4.6:** Components of the toy model of the SPR results, as a function of the AI and genitive argument status. a: The ESR. b: Dominance. c: Coherence (see text for details). d: Approximation of acceptability ratings as the mean of dominance and coherence. e: Approximation of acceptability ratings as the result of taking the square root of the dominance values (solid line) and linear smooths of the square root (broken lines).

place once the whole sentence has been parsed and the rating is to be provided.

The toy model just sketched is based upon a handful of rather simple assumptions, including the concepts of ISP, ESR and dominance, which suffice to jointly regulate the competitive linking process between the co-arguments via a simple mutual suppression mechanism. While



this simple toy model obviously is not enough to predict actual RTs, it suffices to derive the characteristic interaction pattern of the 'AI' and 'Linking' predictors present in the RTs for the 'Genitive' and 'PostGenitive 1' regions, which I have argued to mainly reflect the mobilisation of additional suppression resources in form of the ESR. In addition, the toy model also allows to explain the acceptability ratings on basis of the dominance index, which can be understood as a simple measure of the extent of the eventual prominence of the currently processed genitive argument over its implicit co-argument; one assumption here is that the mapping from dominance to acceptability is not one-to-one, but that dominance still provides the key component from which the degree of acceptability is derived.

Note that many of the assumptions behind the mechanism sketched above can be understood as reflecting properties of the structural architecture of the German NP and the nominal linking mechanisms arising from the grammar of *-ung*-nominalisations. Thus, as discussed in section 2.1, in her discussion of linking patterns found with deverbal, eventive nominalisations, Stiebels (2006) argues that the key properties of these can be traced back to the grammar's need to find a way to meet the challenge of accommodating verbal 'riches' given the structural 'rags' encountered in the nominal linking system. In the German NP, the availability of only one fully-fledged argument position for genitives leads can be interpreted as leading to a competition of the GenO and GenS arguments of deverbal nominalisations of syntactically transitive base verbs. The acceptability rating results of the two experiments conducted so far as well as the RT patterns found in the regions critical for integration of the genitive argument in the current SPR study suggest that GenO's degree of affectedness – and thus its semantic prominence – indeed plays a crucial role in regulating this linking competition.

So far, the semantic feature I have focussed upon in the presentation of the toy model has been AI-induced semantic prominence; in the preceding chapters, I have already discussed a closely related and complementary perspective, which shifts the focus onto the correlates of this kind of semantic prominence, i.e. the prototypicality of GenO, the semantic distance of the co-arguments and thus the degree of semantic transitivity implied by a verb. In fact, thinking of the mechanisms outlined as a consequence of semantic (dis-)similarity of the co-arguments may actually provide a conceptually more interesting foundation: with decreasing prototypicality of GenO and semantic distance between the co-arguments, they become semantically more similar. This semantic approximation of GenO and GenS may be interpreted as the source of the mechanism regulating the allocation of ISP. On this view, it is the increasing semantic similarity between the co-arguments which causes the more equal distribution of the ISP: when the AI is at its minimum, the co-arguments are at or near maximum similarity and the ISP is thus distributed (roughly) equally among them.

Importantly, the mechanism sketched above is also compatible with a number of independently proposed principles of sentence processing, such as the role of prominence information and the Distinctness principle within the eADM. The results of the SPR experiment suggest that prominence is indeed a factor which regulates the argument linking in the current paradigm online and thus provides initial evidence that verbally induced semantic prominence can in principle impact sentence processing in a way similar to that of NP-inherent prominence information, which has largely been the focus of the eADM. In fact, the toy model outlined

can be interpreted as a specialised analogue of the eADM's Compute Prominence step which takes the context-specific conditions encountered in the constructions used in the present experimental paradigm into account.

The current results can also be related to the eADM's principle of Distinctness (see Bornkessel-Schlesewsky & Schlewsky, 2009), which states that one factor determining the ease of the argument linking process is the degree of (dis-)similarity of co-arguments of a verb in all relevant aspects of prominence: when the different event participants are maximally prototypical and thus also dissimilar, the integration proceeds with less processing effort; decreasing prototypicality and distance between the semantic profiles of the co-arguments, on the other hand, incurs processing cost, since the less distinct GRs enter into competition with one another (see section 3.2.4.1). In the present paradigm, the impact of Distinctness depends upon the type of argument processed: from the perspective of GenOs, the effect of decreasing prototypicality – and increasing semantic similarity to GenS – is indeed detrimental, since it correlates with a more equal distribution of the ISP, and thus with increased competition by the GenS co-argument and allocation of additional ESR. For a GenS argument, however, decreasing distinctness from GenO entails a higher share of the ISP and thus increased competitiveness and less ESR. Note that these beneficial effects of similarity arising with GenSs are simply the consequence of the competition mechanism proposed above, which reflects the specific linking conditions encountered with *-ung*-nominalisations – in general, I believe that the results are in line with research on the impact of similarity-based interference in sentence processing (see section 3.2.5) and the principle of Distinctness, suggesting that Distinctness may also apply to verb-induced semantic properties of argument NPs, thus potentially widening the scope of the principle as presented in the eADM.

#### 4.1.8 Summary

The acceptability ratings of the present SPR experiment reproduced the results of the prior rating study, yielding additional support of the role of the AI in regulating argument linking processes with German *-ung*-nominalisations. The analysis of the RT data for the 'Nominalisation' and 'Genitive Article' regions failed to find evidence for processing cost for predicates with a higher AI, a null-result which may have been the consequence of counteracting factors which were not controlled for in the current experiment. While the study thus failed to find processing correlates of predicate-related lexical-semantic complexity (research question one), the current RT results support the hypothesised role of predicate-induced semantic prominence (affectedness) in resolving argument linking conflicts with *-ung*-nominalisations in online sentence processing (research question two). I outlined a toy model of the underlying competition between the GenO and GenS co-arguments which conceptualises this process as a mutual suppression mechanism primarily regulated by GenO's semantic prominence, as defined by its degree of affectedness. This toy model explains the RT patterns as the consequence of the parser's mobilisation of additional suppression resources (the ESR) in its attempt to integrate any genitive argument encountered in accordance with the respective target linking. In addition, it approximates the acceptance patterns via the eventual dom-

inance of the argument to be integrated over its implicit co-argument, a notion which is a close relative to and direct derivative of GenO's semantic prominence. The results are thus generally in line with the role of prominence information and the principle of Distinctness as posited in the eADM (Bornkessel & Schlesewsky, 2006; Bornkessel-Schlesewsky & Schlesewsky, 2009); the ERP experiment presented in the subsequent section investigates the electrophysiological correlates of the underlying processes in detail.

## 4.2 ERP experiment

### 4.2.1 Materials

The stimuli used for the ERP experiment were based upon the stimulus material used for the SPR experiment and constructed using the same procedure (see section 4.1.1 for details). However, the self-paced reading stimuli were adapted in a number of ways: to enhance the degree of naturalness of the stimuli, some of the stories were revised. Further, to somewhat increase the density of nominalisations at the very low end of the AI scale and thus obtain a more balanced data set, an additional pretest was conducted to collect AI ratings for a number of additional nominalisations. To this end, an abridged version of the AI-rating pretest described in section 2.7.2.1 was constructed: 10 previously unused *-ung* nominalisations whose syntactically transitive base verbs only imply none to moderate obligatory change for their respective direct object participant (which were selected following the criteria listed in section 2.7.2.1) were mixed with 30 further nominalisations already used in the previous experiments; these were already associated with an AI value and represented the whole range of the AI scale, resulting in a total of 40 *-ung* nominalisations. Using these nominalisations' base verbs, two different short SVO sentences were constructed as described in the 'Materials' section of the first AI-rating pretest; this yielded two experimental lists, each of which featured all 40 nominalisations' base verbs, but with different lexical fillers for the subjects and direct objects. The two lists were used to construct an abridged spreadsheet version of the first AI-rating pretest, which was conducted as an internet rating study; instructions for the ratings task were provided as detailed in section 2.7.2.1 and adapted to fit the spreadsheet version. Ratings were again provided on an integer scale ranging from 1–7, with 1 indicating that the relevant event participant definitely did not undergo any kind of change (or that it is left completely open whether it did) and 7 indicating that it definitely underwent a specific change. 11 participants (mean age 41.9; 9 female, 2 male) completed the questionnaire and were assigned to one of the two experimental lists on an alternating basis. The order of stimulus sentences within a list was randomised separately for each participant.

From the 10 new nominalisations, 5 were selected for the construction of stimuli for the ERP experiment. Since 5 other nominalisations used in the self-paced reading experiment were excluded, the resulting number of nominalisations was again 84, as in the SPR study. These 84 nominalisations were used to construct the stimulus material in the same way as described in section 4.1.1 for the SPR experiment, resulting in two experimental lists with

168 stories each. Each of the lists was divided into two blocks, such that each nominalisation only occurred once within a block. The order of presentation of items within a block was randomised separately for each participant; participants were assigned to one of the two lists on an alternating basis and the order of presentation of blocks within a list alternated as well. Information on the nominalisations used is provided in table A.1 and the complete stimulus material is listed in Appendix D.

#### **4.2.2 Procedure**

The experiment took place in a dimly lit and sound-attenuated chamber; participants faced a computer screen, which was placed approximately at a distance of 1 m from the subjects. The experimental stimuli were delivered using Presentation software (Neurobehavioral Systems, USA) and presented in black letters (Arial font; font size 22) against a light-grey background. Each trial started with the presentation of the complete context sentence, which participants read without a time limit. Upon having read the context sentence, participants pressed the enter key, which started presentation of the continuation sentence in a word-by-word, rapid serial visual presentation (RSVP) mode. Before the first word of a continuation sentence, a fixation cross was presented at the centre of the screen for 800 ms, followed by presentation of a blank screen for 200 ms. Subsequently, each word of the continuation sentence was presented at the centre of the screen for 250 ms, followed by a blank screen lasting for 350 ms (SOA = 600 ms; ISI = 350 ms). Experimental items were preceded by 8 practice items.

After each story, participants rated the acceptability of the second sentence as a coherent continuation of the first sentence on an integer scale ranging from 1–6 by pressing the respective button on the keyboard, with '1' indicating complete unacceptability as a coherent continuation and '6' straightforward acceptability; instructions for the rating task were given as described for the SPR experiment in section 4.1.2 above. Participants were again instructed to provide their ratings swiftly and response time for the ratings was limited to 6 seconds; when participants exceeded this time limit, they received feedback about being too slow. A subset of 36 of the stories was followed by short 'Yes'/'No' comprehension questions to make sure participants were paying attention to the contents of the sentences. These targeted different aspects of the sentence pairs presented. Participants received feedback in case of incorrect responses to these questions.

After every 12 items, participants were prompted to make a short break. One experimental session lasted for approximately 1½ – 2 hours, including preparation of the EEG electrodes.

#### **4.2.3 EEG recording and preprocessing**

The EEG was recorded from 61 Ag/AgCl electrodes mounted in an electrode cap (Easy-Cap™) and one further electrode was placed at the left mastoid. In addition, the horizontal and vertical electrooculograms were monitored using two electrodes at the outer canthi of the eyes and one electrode below the left eye. Offline, the two electrodes placed at the

outer canthi of the eyes were reused as electrodes F9 and F10; thus, the final data set comprised the following 63 EEG electrodes: Fp1, Fpz, Fp2, AF7, AF3, AFz, AF4, AF8, F9, F7, F5, F3, F1, Fz, F2, F4, F6, F8, F10, FT9, FC5, FC3, FC1, FCz, FC2, FC4, FC6, FT10, T7, C5, C3, Cz, C4, C6, T8, TP9, TP7, CP3, CP1, CPz, CP2, CP4, TP8, TP10, P9, P5, P3, P1, Pz, P2, P4, P6, P10, PO9, PO7, PO3, POz, PO4, PO8, PO10, O1, Oz, O2. During recording, all channels were referenced to the left mastoid; electrode impedance was kept below 10 k $\Omega$ ; signals were sampled at 500 Hz and band pass filtered at 0.016–70 Hz. All further preprocessing of the recorded EEG signals was performed using BrainVision Analyzer 2.0.2 (Brain Products GmbH, Germany) and BESA 6.0 (BESA GmbH, Germany) software; offline, the continuous EEG was re-referenced to an average reference montage and corrected for blinks using Surrogate Multiple Source Eye Correction (Ille, Berg, & Scherg, 2002). After application of a 30-Hz low pass filter and downsampling of the EEG to 250 Hz, the continuous EEG was segmented into 1400 ms epochs around the onsets of the critical words (nominalisations and genitives), starting 200 ms before word onset. Artifact-contaminated segments were discarded using a fully automatic procedure. The resulting single-trial level segments were referred to a 200 ms pre-stimulus baseline and exported for further statistical analysis in R (R Development Core Team, 2013).

#### **4.2.4 Hypotheses**

##### **Nominalisation segment**

Given the relative scarcity of relevant electrophysiological data on the effects of lexical-semantic complexity (see section 3.2.1) as well as the failure to find a behavioural correlate of affectedness in the SPR experiment presented above, the present investigation of the electrophysiological correlates of the AI remains largely exploratory. Considering the findings of Steinhauer et al. (2001) and Malaia et al. (2009), one may predict an increase in the AI covariate to correlate with a more negative-going ERP over anterior scalp regions. Such a prediction may be considered questionable given the failure of Brennan and Pykkänen (2010) to reproduce behavioural effects of verbal lexical-semantic complexity using MEG and the differences between the current experiment and the studies of Steinhauer et al. (complexity of verbal vs. nominal semantic structure) and Malaia et al. (e.g. local vs. non-local relationships). Note that the novel, non-standard analysis approach to the ERP data described below in detail compensates somewhat for the partially exploratory nature of the current experiment, by allowing to analyse the full spatiotemporal extent of the segments, rather than concentrating on pre-defined windows or regions of interest.

##### **Genitive segment**

The SPR experiment produced evidence for an interaction between the 'Linking' and 'AI' variables as the parser attempted to integrate the genitive argument with the respective nominalisation. I conjectured that this interaction pattern may reflect mobilisation of ESR to maximise

the chances of suppressing the encountered genitive's (implicit) co-argument. Crucially, the ESR directly depends upon the AI associated with a nominalisation's GenO argument, i.e. its predicate-induced semantic prominence, or – alternatively – the prototypicality of GenO and the semantic distance of the co-arguments. In the discussion of the findings of the SPR experiment, I pointed out that they are in principle in line with principles of sentence processing posited within the eADM (Bornkessel & Schlesewsky, 2006; Bornkessel-Schlesewsky & Schlesewsky, 2009), such as the impact of prominence information and the Distinctness principle, and that the mechanism sketched may be understood as a contextually specialised variant of the eADM's Compute Prominence step (see section 4.1.7).

Within the eADM, the main electrophysiological correlate of the Compute Prominence (as well as of the Compute Linking step) postulated is the N400, as outlined in section 3.2.4. Thus, the following hypothesis tests if the linking conflicts to be resolved in the present experiment fall within the scope of the N400:

- Hypothesis 1: The N400 should be modulated in a way which is in accordance with the interaction pattern found in the SPR experiment, which was tentatively interpreted as reflecting mobilisation of ESR: in the GenO condition, an increase of the AI should correlate significantly with a more positive N400, while it should correlate with a more negative-going N400 in the GenS condition; at the low end of the AI scale, the N400 potential should be about equal across the GenO and GenS conditions.

Hypothesis 1 for the genitive segment is largely based upon the model-theoretic assumptions of the eADM. However, given the differences between the eADM's view of prominence in argument linking and the current focus, which I discussed in section 3.2.4.3, the hypothesis projects somewhat beyond the model's domain: while the conception of the N400 as a marker of regulating conflicts related to the (semantic) prominence of arguments is mainly built upon NP-inherent characteristics (e.g. animacy and definiteness) in the eADM, the notion of semantic argument prominence investigated in the present work depends upon characteristics induced by the predicate. Hypothesis 1 is thus a test of the scope of the N400 as a *general* processing correlate related to different types of argument prominence. Note, however, that the current experiment does not include constructions which test effects related to the kind of NP-inherent argument prominence the eADM focuses upon.

Also note that it is not clear whether the functional interpretation of the N400 matches well with the characterisation of the ESR, which I proposed to reflect the allocation of cognitive resources for suppression of the co-argument; in neural terms, this can be most directly translated into activation of inhibitory resources. To the best of my knowledge, at least in the interpretation of standard (i.e. plausibility) N400 effects, inhibitory activity is usually not assumed to be a prominent component underlying the N400. Nevertheless, hypothesis 1 provides a fruitful starting point for the interpretation of the results of the present experiment.

Within the eADM, the N400 may be followed by a number of functionally distinct late positivities, one of them indexing a well-formedness check monitoring for ill-formed or marked structures, which is assumed to occur as part of a biphasic N400-late positivity pattern (see

section 3.2.4.2 and Haupt et al., 2008). Given the need to evaluate the outcome of the competition process dominated by the allocation of the ESR within the toy model proposed as well as the discrepancy between the RT patterns and the acceptability ratings provided (as discussed in section 4.1.7), hypothesis 2 complements the prediction about the N400:

- Hypothesis 2: The predicted N400 pattern should be followed by a late posterior positivity. The activation pattern should reflect the degree of coherence attributed to the predicate-argument complex in the continuation sentence given the target linking configuration established within the context sentence, with one possible pattern illustrated by figure 4.6c: if the late positivity is modulated in a way in line with this illustration, activation should be about equal for GenO continuations, regardless of the AI, as well as for low-AI GenS continuations. The late positivity should thus show an interaction of the 'AI' and 'Linking' variables; the simple slope of 'AI' within the GenS condition should show a significant positive correlation; in the GenO condition, on the other hand, the 'AI' variable should not modulate the late positivity significantly.

#### **4.2.5 Participants**

28 native German speakers took part in the experiment; the data of one of these had to be excluded from analysis due to excessive EEG artifacts and one more subject was excluded from analysis due to major self-reported difficulties with understanding and carrying out the rating task. The remaining 26 subjects had a mean age of 29.2 (range 19–30; 19 female, 7 male); all were right-handed according to self-reports and the Edinburgh Inventory (Oldfield, 1971) and had normal or corrected-to-normal vision. In return for participation in the experiment, they received either course credits or a monetary compensation of € 18,-. Prior to the start of the experiment, all participants provided written and informed consent.

#### **4.2.6 Data analysis**

##### **Rating data**

The rating data were analysed using LMMs and complementary CLMMs, following the same procedure applied for analysis of the rating data of the SPR experiment (see 4.1.5 above). Since the results of the CLMMs again converged well with those of the LMMs, they are not discussed further; the CLMMs' summary tables are provided in Appendix D.

##### **ERP data**

The ERP data were analysed using an innovative approach which is not part of the standard ERP analysis methods toolbox, applying the wavelet-based functional mixed model (WFMM) of Morris and Carroll (2006) as analytical engine. As pointed out in the Introduction, the two main motivations for taking this approach were to allow for consistence in the experimental

design throughout the different experiments by retaining the AI as a numerical, continuous covariate, on the one hand, and to compensate for the partially exploratory character of the present ERP study, on the other hand. These two issues are solved by using a method which, among other things, allows to model ERP data without prior definition of temporal windows or spatial regions of interest and to account for effects of continuous predictors in a natural way.

The WFMM extends LMMs for standard scalar (i.e. 'point') data for the analysis of complex curve data within the framework of functional data analysis (FDA); FDA provides specialised statistical tools for analysing curve data, focussing on curves as single entities, rather than a collection of loosely related data points (see Levitin, Nuzzo, Vines, & Ramsay, 2007; Ramsay & Silverman, 2005, for introductions to FDA). One of FDA's foremost goals is to capture the correlational structure present within curves on the respective domain appropriately; often, the relevant continuum is time or space, but it may as well be any other "continuous psychophysical space (such as vowel space, musical pitch space, or colour space)" (Levitin et al., 2007, p. 136). Many applications involve 1d-curve<sup>1</sup> data, but the FDA repertoire also comprises tools for the analysis of multidimensional curve data, such as 2d-images or 3d-volumes (e.g. Morris et al., 2011).

Note that chapter 6 in the second part of this thesis provides an in-depth discussion of standard LMMs, challenges of analysing curve data, the FDA framework, conceptual and technical details of the WFMM and related topics (such as properties of wavelets and computational issues and limitations); the following outline of the WFMM will thus focus on the most important technical and conceptual points and the interested reader is referred to chapter 6 for further discussion.

Assuming that all  $N$  curves are sampled on the same fine grid  $t$  of length  $T$ , the WFMM can be expressed as in equation 4.1 (Morris & Carroll, 2006, p. 184):

$$Y = XB + ZU + E; \quad (4.1)$$

here,  $Y$  represents an  $N \times T$  data matrix of the  $N$  curves sampled on  $t$ ;  $B$  is a  $p \times T$  matrix of functional fixed effects;  $U$  an  $m \times T$  matrix of functional random effects and  $E$  the residual error matrix of size  $N \times T$ .  $X$  is the  $N \times p$  design matrix for the fixed effects curves and  $Z$  represents the  $N \times m$  design matrix for the random effects curves. The random effects  $U$  are assumed to follow a matrix normal distribution  $U \sim \mathcal{MN}(P, Q)$ , with  $P$  the  $m \times m$  between-row covariance matrix and  $Q$  the within-row  $T \times T$  covariance matrix;  $E$  is assumed to be  $\mathcal{MN}(R, S)$ , with  $R$  and  $S$  of dimensions  $N \times N$  and  $T \times T$ , respectively. Note that the model formulated in equation 4.1 represents the functional analogue of the linear mixed effects model often applied in the analysis of scalar data.

---

<sup>1</sup>Of course a curve involves data along two axes, such as the single-channel ERP across time. However, one can distinguish between the dimensionality of the data itself and the domain on which it is defined. In case of single-channel ERP data, the data is one-dimensional (the ERP in  $\mu V$ ) and its domain is one-dimensional as well (time). The dimensionality indicated here refers to the dimensionality of the domain only. This is a common way of classifying functional data in FDA in terms of their dimensionality (see Morris, Baladandayuthapani, Herrick, Sanna, & Gutstein, 2011, for example).



Fitting a WFMM follows a three-step procedure, beginning with transforming the  $N$  sampled curves forming the data-space matrix  $Y$  to the wavelet space by means of a discrete wavelet transform (DWT), yielding the wavelet space data matrix  $D$ , whose columns holding the wavelet coefficients are double-indexed by scale  $j$  and location (translation)  $k$ . Subsequently, the resulting wavelet coefficients are taken as the dependent variables for separate models. Parameter estimation is performed within the Bayesian framework (for introductions to the Bayesian approach to data analysis, see Kruschke, 2010a, 2010b, 2010c) via Markov chain Monte Carlo (MCMC) sampling; the result is a set of wavelet-space MCMC samples for a number of functional parameters (e.g. fixed and random effects curves; variances). In a last step, these are back-projected to the data space via the inverse discrete wavelet transform (IDWT), allowing for straightforward interpretation of the estimates. Any kind of Bayesian inference can subsequently be performed on these data space functional posteriors (e.g. constructing pointwise or simultaneous credible intervals and computing posterior probabilities).

Performing the model fitting in the wavelet space, rather than in the original data space, brings along a number of benefits: thus, justified by the whitening property of wavelets, the model's covariance matrices are assumed to be diagonal, thus drastically reducing the number of covariance parameters from  $T(T + 1)/2$  to  $T$ , which enhances computational feasibility. In addition, wavelets are extremely flexible and can be used to represent a wide number of different kinds of curves, whether smooth, spiky and irregular, or both. Importantly, in the spirit of FDA, the properties of wavelets and the architecture of the WFMM enable to borrow strength for estimation from adjacent and nearby time points, resulting in more efficient estimates and inference.

The output of the WFMM includes data space posteriors for the specified fixed and random effects curves and the respective variance curves. The fixed effects posteriors can be used for inference about effects in a time-varying manner, i.e. effects can be assessed at the original temporal resolution without the need to pre-define temporal windows of interest. Note that the fixed effects curves are smoothed during the fitting procedure in the wavelet space by using a spike-and-slab mixture prior for them, which has the effect of shrinking small wavelet coefficients non-linearly further towards zero. This shrinking results in adaptively regularised estimates for the fixed effects curves, preserving prominent local peaks. Thus, the degree of smoothness can vary flexibly along  $t$  as indicated by the underlying data. Since different fixed effect functions are regularised separately, the global and local smoothness properties can vary between the different fixed effects curves of a model.

The WFMM further allows to benefit from the advantages of LMMs: just as a regular LMM, the WFMM can accommodate categorical or continuous covariates. In addition, it can deal with multiple layers of variability in one and the same model, allowing to jointly include random effects (random intercepts and/or slopes) for subjects and items, or other relevant experimental units. The input to the WFMM consists of the single-trial level data, allowing to make use of the full information contained in the data and to model inter-trial dependencies, for example, without prior averaging within subjects and conditions. The WFMM thus accommodates a wide range of experimental designs. Note that the current version of the

WFMM core software does not allow to include parameters to model correlations between random effects (as possible with *lme4*'s `lmer` function, for example).

Previous work using the WFMM for the analysis of ERP data includes Davidson (2009), Davidson, Hanulíková, and Indefrey (2012) and Steen (2010), all of which applied the WFMM to single-channel data or to data representing (multiple) spatial regions of interest (ROIs). The current analysis goes beyond these studies by fitting separate WFMMs to each electrode's data (i.e. a total of 63 WFMMs per segment), thus modelling the ERP data without pre-specifying temporal windows or spatial regions of interest. All analyses were performed using the *stepmom* (spatiotemporal electrophysiological model maps) package for the statistical environment R, which is based upon the *wrapfmm* package; both of these packages are presented in separate chapters in part 2 of this thesis (chapters 8 and 7, respectively).

The two models for the nominalisation and genitive segments were fit using *stepmom*'s `wfmm` method; the respective model specifications in `lmer` notation are provided in (126). All independent variables were again centred on their mean (as indicated by the prefix 'c. '); continuous numerical covariates were additionally standardised by dividing by 2 SDs (as recommended by Gelman, 2008, hence the prefix 's.').

(126) a. *Nominalisation segment*:

$$\text{ERP} \sim \text{s.AI} + \text{c.Embedding} + \text{s.List.Pos} + \\ \text{s.Word.Length} + \text{s.log.Frequency} + \\ (1 \mid \text{Subject}) + (0 + \text{s.AI} \mid \text{Subject})$$

b. *Genitive segment*:

$$\text{ERP} \sim \text{c.Linking} * \text{s.AI} + \text{c.Embedding} + \text{s.List.Pos} + \\ \text{s.Word.Length} + \text{s.log.Frequency} + \\ (1 \mid \text{Subject}) + (0 + \text{c.Linking} \mid \text{Subject}) + \\ (0 + \text{s.AI} \mid \text{Subject}) + (0 + \text{c.Linking:s.AI} \mid \text{Subject})$$

The models for both segments comprised 's.AI' as fixed effect plus by-subject random intercepts and random slopes for 's.AI'; 'c.Embedding', 's.List.Pos', 's.Word.Length' and 's.log.Frequency' were entered as control covariates and are interpreted as described in section 4.1.5 above; lemma frequencies were again obtained from the *dlxDB* database (Heister et al., 2011). In addition to these predictors, the genitive segment's model contained 'c.Linking' as well as the 'c.Linking:s.AI' interaction as fixed effects as well as by-subject random slopes. Note that both models lack random effects for items; depending upon the dimensionality of the data and the complexity of the model, running a WFMM may be computationally challenging and the addition of random effects for items (taking 'Story' and/or 'Word' as relevant grouping factors) would have increased the computation time for each of the 63 single electrode's WFMMs (for each of the two segments analysed) significantly. Considering that computation of a single electrode's model already took several hours without item-related random effects, the models were restricted to subject-related random effects. The complementary standard LMM analyses conducted (see below) are intended to compensate for this shortcoming. The number of trials which entered the analysis was 4255 (97.4% of the

original number of trials) for the nominalisation segment and 4262 (97.6% of the original number of trials) for the genitive segment.

The WFMMs were run using 'db4' wavelets of the Daubechies wavelet family, whose choice was motivated by previous studies (e.g. Davidson, 2009)<sup>2</sup>, with 8 levels of decomposition and a periodic boundary correction. The WFMM software offers an option to perform joint wavelet compression in a way which excludes wavelet coefficients insignificant across all curves while preserving a predetermined amount of energy, which allows to reduce computation time without loss of relevant information. For the current analyses, WFMMs were run preserving 90% of the original amount of energy across curves, which reduced the number of coefficients to be modelled to about 35–55% of the original number (depending on the electrode). 10000 MCMC samples were obtained for each model; the first 1000 of these were discarded as burn-in and the remaining 9000 were thinned by keeping every third sample, i.e. the final number of MCMC samples was 3000. Convergence of the sampling chains was assessed informally with the help of *stepmom's* `chain_plot` function for a number of different electrodes, time points and predictors; inspection of the trace and density plots generally suggested good convergence of the sampling chains.

Planned post-hoc assessments again included testing for simple slope effects of the 's.A1' predictor for the genitive segment. Estimates for these were computed from the original fixed effects posteriors using the `follow_up` function.

When analysing ERP data with *stepmom's* `wfmm` method, a number of default inferences are computed from the fixed effects posteriors, including a number of different Bayesian credible intervals. When presenting estimates for selected fixed effect curves and electrodes across time using the `summary` method, two different types of credible intervals are presented, one being a 95% pointwise credible interval, which is a quantile-based credible interval computed independently at each time point. The other type of credible interval represents a 95% functional simultaneous credible band (SCB) which takes multiplicity across the grid  $t$  into account. These SCBs are more conservative than the pointwise intervals; they correspond to the 'type I' SCBs offered by the `summary` method and are computed as presented in section 3 of Krivobokova, Kneib, and Claeskens (2010)<sup>3</sup>.

In addition, inference about the significance of fixed effects can be based upon different types of posterior Bayesian probabilities; each type of credible interval is accompanied by a corresponding probability type, which is designed to flag the same set of time points as significant as the respective credible interval (or a very close approximation to this set). The posterior probabilities used for assessment of significance here are pointwise probabilities (flagging the same set of grid points as significant as the pointwise credible intervals) and a novel type of Bayesian 'simultaneous' probabilities, the MULTIBONDS (**M**ultiplicity-**I**nduced **B**and **O**f **N**o **D**ifference **S**cores), which are an original contribution of this thesis and are intended to flag the same set of grid points as the type I SCBs (for a similar approach to multiplicity-adjusted posterior probabilities, see Meyer, Coull, Versace, Cinciripini, & Morris, 2015). Details about the computation and interpretation of the MULTIBONDS and all other types of credible in-

---

<sup>2</sup>Other types of wavelets may well be a better fit for typical ERP data (e.g. wavelets of the 'Symlet' family).

<sup>3</sup>See also Steen (2010) for an application of these SCBs to WFMM posteriors.

tervals and posterior probabilities are provided in section 7.7 in part two of the thesis. One of the main advantages of the MULTIBONDS is that they account for multiplicity without requiring the specification of a rather arbitrarily determined minimum effect size to compute Bayesian false discovery rates, as in Morris, Brown, Herrick, Baggerly, and Coombes (2008) and Davidson (2009), for example; rather, MULTIBONDS are computed from the posterior on a purely objective basis without prior specification of minimum effect sizes.

*stepmom's mom (model map)* function allows to plot various types of posterior quantities as scalp maps for arbitrary time windows, in the spirit of a number of recent studies such as Hauk, Davis, Ford, Pulvermüller, and Marslen-Wilson (2006), Hauk, Pulvermüller, Ford, Marslen-Wilson, and Davis (2009) and Amsel (2011b). In the following presentation of the results, plots showing (interpolated) scalp maps of the posterior mean estimate of selected predictors will be presented together with maps of the respective pointwise posterior probabilities and MULTIBONDS; the MULTIBONDS were computed using 95% type I SCBs and the threshold for indicating significance is set to  $\alpha_{MULTIBONDS} = 1 - prob_{SCB} = 0.05$ , which is also the threshold for the pointwise probabilities. The values within the respective model maps represent the mean of each posterior quantity within the indicated time windows.

Note that the pointwise probabilities/credible intervals and multiplicity-adjusted SCBs/MULTIBONDS are presented together intentionally: while it is desirable to provide conservative assessments of the credibility of posterior estimates, the functional nature of the WFMM and the adaptive smoothing performed for the posterior fixed effect curves should already provide a certain degree of protection against false discoveries induced by multiple comparisons across the grid  $t$  (Jeffrey S. Morris, personal communication). Since an in-depth assessment of the degree of conservativity of the SCBs/MULTIBONDS under different conditions is currently lacking, pointwise and multiplicity-adjusted posterior quantities are presented jointly; where both probability types agree upon the significance of a specific cluster, the respective effect can be considered sound; where these two diverge, the respective effect should be interpreted with more caution. Note that clusters identified on basis of simultaneous inference will necessarily be shorter than their counterparts based upon pointwise posterior quantities; also, sometimes a single cluster identified by pointwise probabilities/credible intervals may be split into two or more clusters when inference is performed using simultaneous quantities.

Relevant significant spatiotemporal clusters were identified in the following way: first, table summaries providing information about significant temporal clusters (i.e. consecutive significant time points) of the interest predictors (i.e. 's.A1' for the nominalisation segment; additionally, 'c.Linking' and the 'c.Linking:s.A1' interaction for the genitive segment) were obtained using *stepmom's summary* function, based upon the less conservative pointwise probabilities; subsequently, the *ani* function was used to make animated scalp maps of these predictors over time, starting ~20 ms before the start of the earliest cluster thus identified and lasting until the end of each segment, with frames always aggregating over two adjacent time points. The resulting animations thus represent the full spatiotemporal extent and dynamics of the effects at a temporal resolution of 8 ms; they are available as supplementary materials at

<http://doi.org/10.18452/19291>.<sup>4</sup>

The animations and table summaries were then used to identify significant spatiotemporal effect clusters with stable scalp distributions as well as the electrode and point in time at which each effect cluster reached its peak (as defined by peak mean estimate and/or probability). The model maps presented in the results section show the average posterior mean estimate and the corresponding average probabilities covering a representative time span around the respective maxima; they are complemented by curve plots illustrating relevant ERP patterns and WFMM estimates at the respective peak electrodes over time. Effects of the control covariates are not discussed, but animated model maps illustrating their spatiotemporal evolution throughout the segments at a temporal resolution of 8 ms are again available as electronic supplementary materials at the address provided above.

To compensate for some of the practical limitations of the WFMM analysis, such as the lack of item-related random effects in the present WFMMs, *stepmom*'s `lmerMom` function was used to perform complementary analyses and confirm the most relevant findings by modelling single-trial ERP data averaged within pre-defined time windows with standard LMMs using *lme4*'s `lmer` function (Bates et al., 2014) for each of the 63 electrodes separately. The respective time windows were the same as selected for the presentation of the WFMM-based spatiotemporal effect clusters, selected as described above. The fitting of LMMs with complex random effect structures is frequently complicated in practice by convergence issues, which currently cannot be traced back to specific electrodes using the `lmerMom` function, let alone resolved. To reduce the number of convergence problems, the only item-related random effect added was a random intercept for 'Word' (ignoring 'Story' as a grouping factor); also, since inclusion of random correlation parameters for all subject-related random effects led to numerous convergence warnings, they were removed from the models. The resulting model specifications were as given in (126), plus by-Word random intercepts. For these additional analyses, the single trial time-window mean data were trimmed by excluding data points more than 2.5 SDs above/below the mean separately for each participant, electrode and time window.

The results of these LMM analyses can again be presented as scalp model maps showing the estimates of the respective predictors of interest. These are accompanied by  $p$ -value scalp maps demarcating significant regions; the  $p$ -values used for these maps are lower-bound degrees of freedom  $p$ -values obtained using the `pamer.fnc` function of package *LMERConvenienceFunctions* (Tremblay & Ransijn, 2013). The  $p$ -values were corrected using a nearest-neighbour based correction algorithm, which only uses a set of  $k+1$  neighbouring sensors to correct  $p$ -values for each electrode (and time window), rather than the full set of electrodes; the  $p$ -values associated with each electrode's model were corrected using all neighbouring electrodes located within a radius of approximately 3 cm by means of a false-discovery rate procedure<sup>5</sup>. This induces a lenient correction which accounts for different electrode densities varying locally across the scalp. Since the results of the complementary `lmerMom`-based models generally converged well with those of the WFMM-based analysis, the former will

<sup>4</sup>See the 'readme.txt' files available with the supplementary materials for more information.

<sup>5</sup>The function used for this correction procedure is `p.adjustSP` from package *spdep* (Bivand, 2013).

not be discussed explicitly here. The resulting model maps for the most relevant aspects are provided in Appendix D.

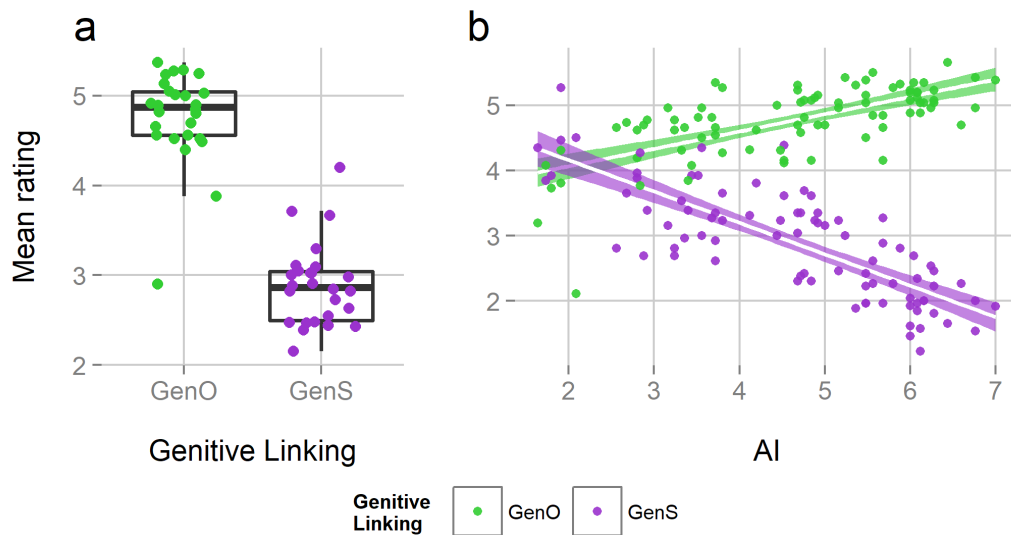
## 4.2.7 Results

### Question response data

On average, participants answered 91.2% of the questions correctly (range: 75% – 100%, SD = 5.3), suggesting that they paid close attention to the contents of the sentences.

### Rating data

The boxplots in figure 4.7a show the mean acceptability scores computed for each subject and 'Linking' level and illustrate that GenOs on average received higher acceptability ratings than GenSs (GenO: mean = 4.77, SD = 0.51; GenS: mean = 2.89, SD = 0.46). Figure 4.7b shows the mean acceptability rating per nominalisation at each level of 'Linking' over the range of 'AI' with separate linear smooths for 'AI' for the 'GenO' and 'GenS' conditions. The linear smooth for the 'GenO' condition suggests good acceptability throughout the range of 'AI' and shows an increase with rising AI values. In the 'GenS' condition, on the other hand, acceptability is as high as for 'GenO' continuations for nominalisations with a minimal AI and decreases as the AI increases, matching the pattern observed with the rating data of the previous experiments.



**Figure 4.7:** Mean acceptability ratings for each 'Linking' level. a: Boxplot of mean acceptability ratings for each subject and 'Linking' level. b: Mean acceptability ratings for each nominalisation at the 'GenS' and 'GenO' levels of 'Linking' plus linear smooths for 'AI' at each level of 'Linking'.

The final specification of the model for the complete data set is provided in (127); the model is summarised in table 4.10.

(127) Rating ~ c.Linking \* c.AI +  
(1 + c.Linking \* c.AI | Subject) + (1 + c.Linking | Story)

The model summarised in table 4.10 shows significant effects of 'c.Linking' ( $B = -1.88$ ,  $SE = 0.16$ ,  $p < 0.001$ ), 'c.AI' ( $B = -0.11$ ,  $SE = 0.02$ ,  $p < 0.001$ ) as well as their interaction 'c.Linking:c.AI' ( $B = -0.74$ ,  $SE = 0.05$ ,  $p < 0.001$ ).

Random effects

Groups	Name	Variance	SD	Corr.
Story	(Intercept)	0.06	0.24	
	c.Linking	0.61	0.78	0.23
Subject	(Intercept)	0.08	0.28	
	c.Linking	0.57	0.75	-0.11
	c.AI	0.00	0.05	0.43 0.74
Residual		1.36	1.17	

Fixed effects

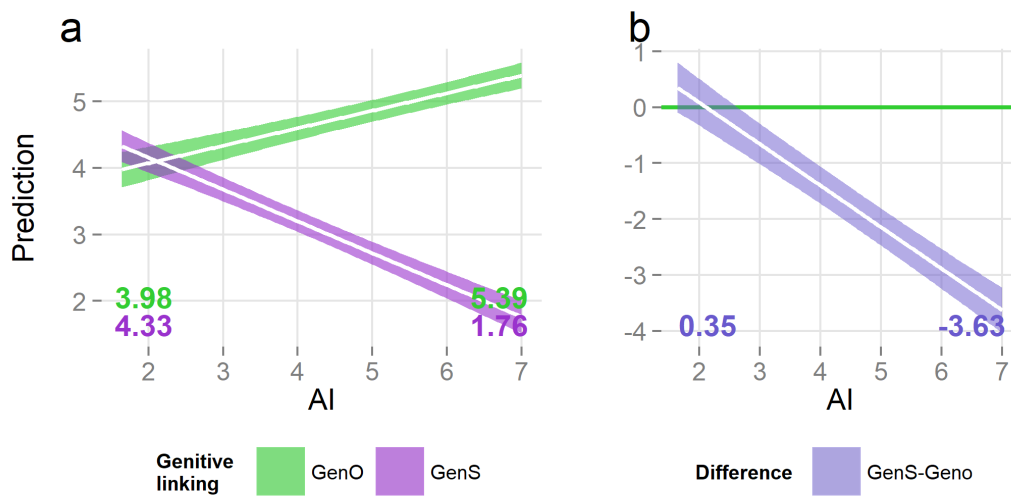
	Estimate	SE	Lower	Upper	t-value	Pr(> t )	
(Intercept)	3.83	0.06	3.71	3.95	62.95	<0.001	***
c.Linking	-1.88	0.16	-2.20	-1.56	-11.49	<0.001	***
c.AI	-0.11	0.02	-0.15	-0.07	-5.22	<0.001	***
c.Linking:c.AI	-0.74	0.05	-0.84	-0.64	-14.76	<0.001	***

**Table 4.10:** Summary tables for the LMM for the acceptability rating data of the ERP experiment.

Figure 4.8 summarises the predictions derived from model (127). Figure 4.8a shows the predicted slopes of 'AI' at the 'GenO' and 'GenS' levels of 'Linking'; it illustrates that the acceptability of a continuation sentence as a coherent continuation was equally high for GenO and GenS occurring with nominalisations associated with a minimal or low AI (estimate at the minimal AI value for GenO: 3.98; for GenS: 4.33). Acceptability decreased with increasing AI in the 'GenS' condition to minimal acceptability at the upper end of the 'AI' scale (estimate at the maximal AI value for GenS: 1.76), while it increased to maximal acceptability in the 'GenO' condition (estimate at the maximal AI value for GenO: 5.39). Figure 4.8b shows the estimates for the corresponding difference between 'GenS' and 'GenO' (i.e. GenS – GenO) across the 'AI' range with the corresponding bootstrap CIs.

Tables 4.11 and 4.12 summarise the post-hoc models fitted to test for simple slope effects of 'c.AI' at each level of 'Linking'; the respective model specifications are provided in (128).

(128) a. *GenO-model*:  
Rating ~ c.AI +  
(1 + c.AI | Subject) + (1 | Story)



**Figure 4.8:** Predictions derived from model (127) with 95% bootstrap CIs. a: Predicted slopes of 'AI' at the 'GenO' and 'GenS' levels of 'Linking'. b: Predictions for GenS - GenO difference across the range of AI.

b. *GenS-model*:  
 Rating ~ c.AI +  
 (1 + c.AI | Subject) + (1 | Story)

The model fitted to the 'GenO' subset (table 4.11) shows a significant effect of 'c.AI' ( $B = 0.26$ ,  $SE = 0.03$ ,  $p < 0.001$ ), i.e. a unit increase on the AI scale is associated with an increase of acceptability of 0.26 on the acceptability scale ranging from 1–6. Table 4.12 shows a significant effect of 'c.AI' in the model fitted to the 'GenS' subset ( $B = -0.48$ ,  $SE = 0.04$ ,  $p < 0.001$ ); thus, in the 'GenS' condition, a unit increase on the AI scale is associated with a decrease of acceptability of  $-0.48$ .

## ERP data

### Nominalisation segment

The WFMM-based analysis of the nominalisation segment (see sections 4.1.1 and 4.2.1) identified rather local effect clusters associated with the 's.AI' predictor. The model maps displayed in figures 4.9a and b illustrate the scalp distributions associated with the 's.AI' predictor in two time windows; in the model maps presented, the posterior mean estimate of any effect is given in the first column and interpolated maps of the pointwise posterior probabilities and MULTIBONDS in the second and third columns; the legend for the probabilities is the one labelled 'p'. Additionally, spatial regions identified as significant by the



Random effects

Groups	Name	Variance	SD	Corr.
Story	(Intercept)	0.18	0.42	
Subject	(Intercept)	0.25	0.50	
	c.AI	0.01	0.08	0.12
Residual		1.21	1.10	

Fixed effects

	Estimate	SE	Lower	Upper	t-value	Pr(> t )	
(Intercept)	4.77	0.11	4.56	4.97	45.28	<0.001	***
c.AI	0.26	0.03	0.20	0.33	8.00	<0.001	***

**Table 4.11:** LMM results for the GenO subset of the ERP acceptability data.

Random effects

Groups	Name	Variance	SD	Corr.
Story	(Intercept)	0.24	0.49	
Subject	(Intercept)	0.20	0.44	
	c.AI	0.02	0.12	0.81
Residual		1.48	1.22	

Fixed effects

	Estimate	SE	Lower	Upper	t-value	Pr(> t )	
(Intercept)	2.89	0.10	2.70	3.08	29.43	<0.001	***
c.AI	-0.48	0.04	-0.56	-0.40	-11.70	<0.001	***

**Table 4.12:** LMM results for the GenS subset of the ERP acceptability data.

more conservative MULTIBONDS are marked by grey and purple lines within the maps of the posterior mean estimates, demarcating areas significant at the 0.05 and 0.001 levels, respectively. While the distributions displayed in figures 4.9a and b are not identical, both are roughly characterised by a frontal positivity and a centroparietal negativity. Within this global distribution, the strongest and most durable effect is a significant left anterior positive shift, which starts at 388 ms post onset and lasts until the end of the segment according to the pointwise probabilities, with a peak effect of 0.74  $\mu$ V at electrode F9 at 1196 ms; this extended effect cluster is broken up into two shorter ones when inference is based upon the MULTIBONDS, with the first of these two starting at 704 ms (see table 4.13 for details). An inspection of the model maps in figure 4.9 shows that this left anterior effect occurs very localised at electrode F9.

The ERP curve plot in the bottom half of figure 4.9 shows the mean ERPs for low, mid and high AI levels at electrode F9 from 250 ms to the end of the segment, with the orange rectangle marking the time windows chosen for representation of the model maps. The pointwise

Predictor	Component	Maximum	Start/End (Length)		Fig.
			BONDS	pointw. <i>p</i>	
s.Al	Left ant. pos.	0.74 $\mu$ V @ F9/1196 ms	704/884 (180 ms)  960/1200 (240 ms)	388/1200 (812 ms)	4.9a/b

**Table 4.13:** Information about the effect clusters associated with the 's.Al' predictor for the nominalisation segment. Summary information is based upon the electrodes at which the effects are maximal. Abbreviations: 'Ant.' = 'anterior'; 'Pos.' = 'positivity'.

95% credible intervals are plotted in light grey, the blue ribbons represent the more conservative 95% SCBs and the posterior mean estimate is marked by the white line inside the pointwise credible intervals. The significance of an estimate at a given time point according to either of these two credible intervals can be derived by checking whether they exclude zero (dark grey line) or not. With relatively weak effects where the credible intervals just cross the zero line and remain close to it, visual identification is not always easy, however. Note that the colour bar at the bottom of the curve plots represents the posterior probabilities of the MULTIBONDS type, which are also plotted in the rightmost column of the model maps.

The two representative model maps for this effect displayed show that during the first phase the left anterior positivity occurs with an equally localised negativity at left posterior electrode P5 (see figure 4.9a) and with a right central negativity during the later time window most prominent at electrode C4 (see figure 4.9b). Note that these negativities are small effects (with maxima of -0.35  $\mu$ V and -0.39  $\mu$ V, respectively).

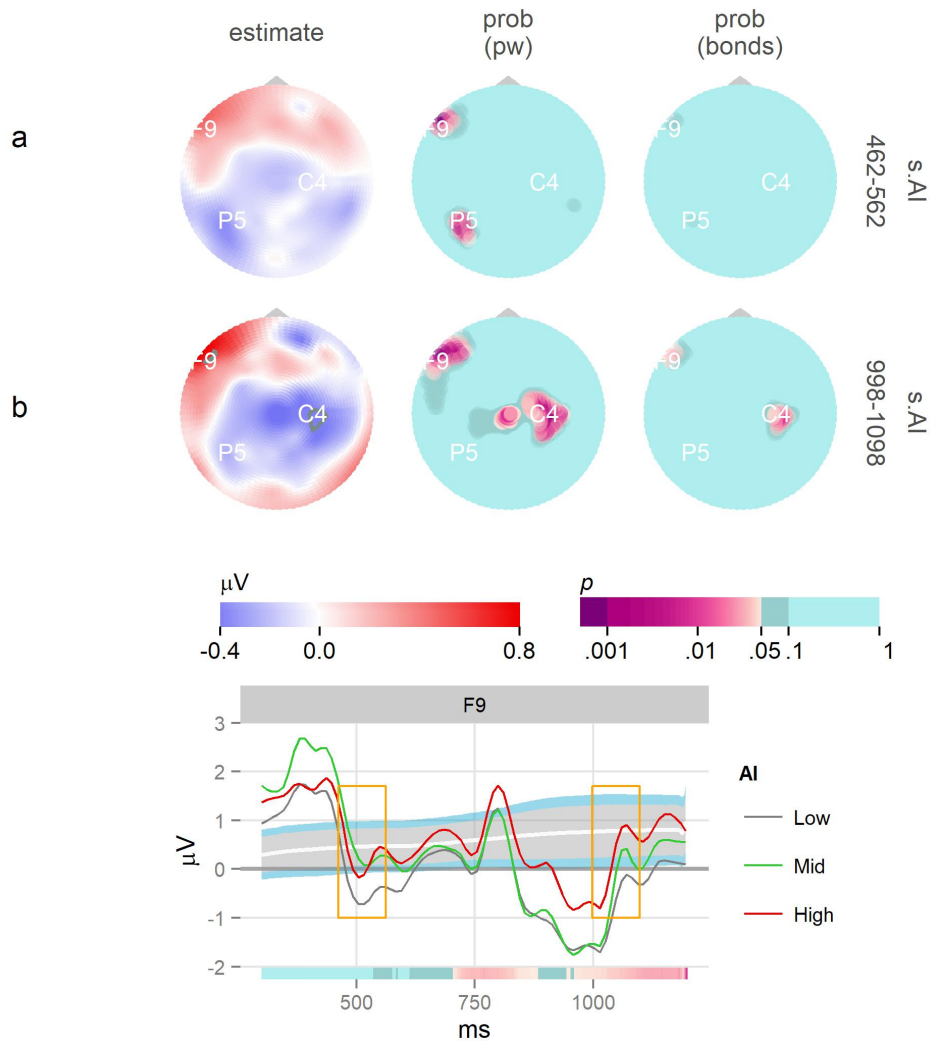
## Genitive segment

Table 4.14 lists the effect clusters identified in the WFMM-based analysis of the genitive segment (see sections 4.1.1 and 4.2.1) along with their details. The effect clusters associated with the 'c.Linking' predictor occur within two phases with distinct global distributions of the mean estimates: the first is shown in figures 4.10a/b and is characterised by a positivity along most of the midline, with a maximum over anterior midline electrodes. This distribution is followed by a distinct one illustrated by figure 4.10c, which is dominated by a positivity with a broad and symmetric centroparietal focus. Within the first of these two phases, 'c.Linking' is associated with an effect cluster over anterior midline electrodes (see row 1 of table 4.14): this anterior midline positivity starts at 368<sub>pw</sub>/ 432<sub>BONDS</sub> ms and lasts until 916<sub>pw</sub>/ 860<sub>BONDS</sub> ms, reaching its maximum effect of 1.12  $\mu$ V at 696 ms at electrode AFz. Figure 4.10b shows the means of the posterior quantities for a 100 ms window around the effect maximum; the map also shows a polarity inversion with a right frontotemporal focus.

This anterior positivity is accompanied by a somewhat weaker posterior positivity with a left-hemisphere focus over electrode P5; this effect starts briefly after the anterior positivity at 424<sub>pw</sub>/ 492<sub>BONDS</sub> ms and lasts until 900<sub>pw</sub>/ 668<sub>BONDS</sub> ms, reaching its maximum effect of

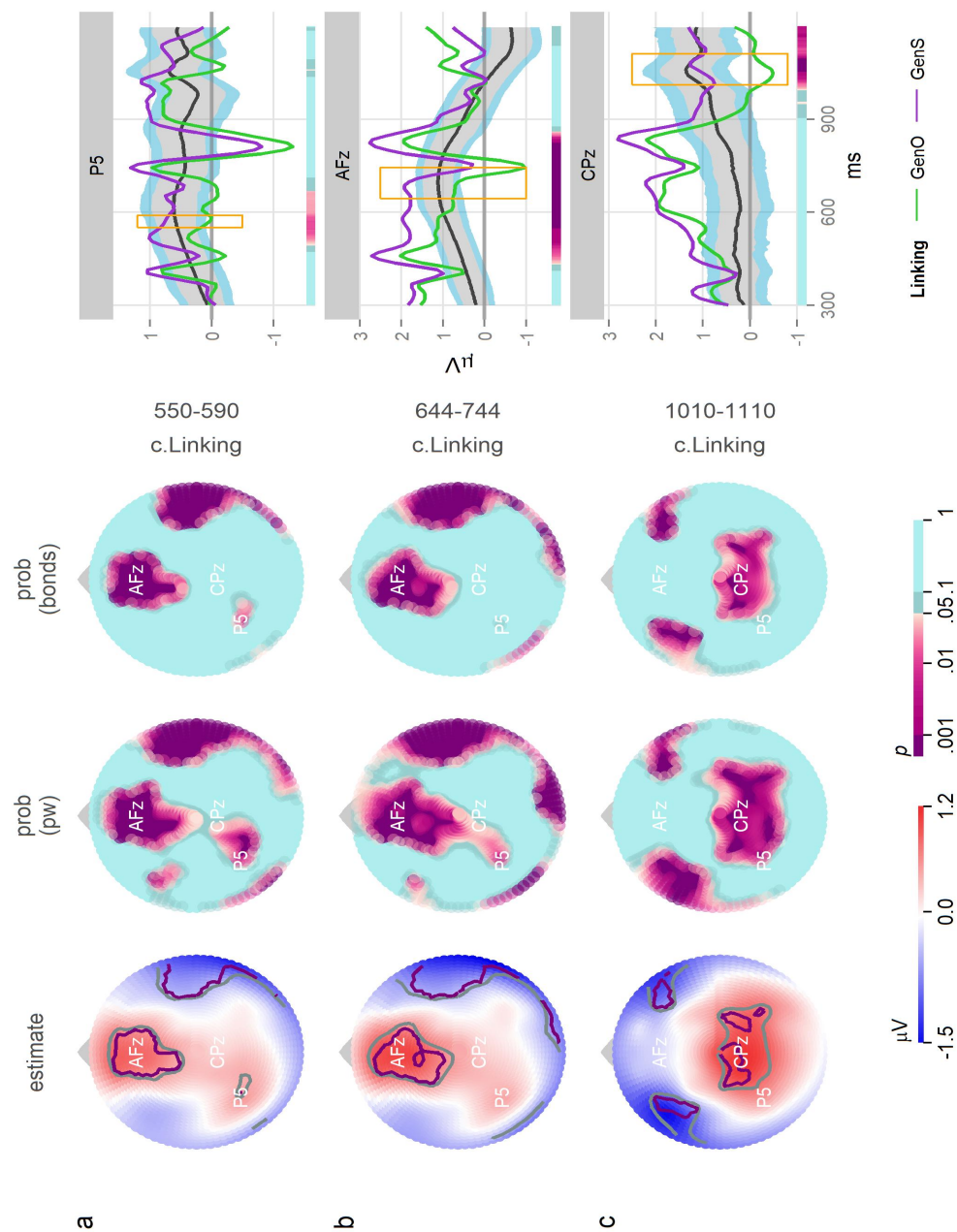
Predictor	Component	Maximum	Start/End (Length)		Fig.
			BONDS	pointw. <i>p</i>	
c.Linking	Ant. midl. pos.	1.12 $\mu$ V @ AFz/696 ms	432/860 (428 ms)	368/916 (548 ms)	4.10b
	Left post. pos.	0.61 $\mu$ V @ P5/620 ms	492/668 (176 ms)	424/900 (476 ms)	4.10a
	Post. midl. pos.	1.36 $\mu$ V @ CPz/1060 ms	992/1200 (208 ms)	800/1200 (400 ms)	4.10c
s.Al	Left post. pos.	0.58 $\mu$ V @ P5/640 ms	444/824 (380 ms)	364/860 (496 ms)	4.11
c.Linking:s.Al	Ant. midl. pos.	1.75 $\mu$ V @ Fz/548 ms	512/660 (148 ms)	492/828 (336 ms)	4.12a
			676/776 (100 ms)		
	Left post. pos.	0.54 $\mu$ V @ P5/740 ms	n.s.	336/824 (488 ms)	4.12a
	Right post. pos.	1.36 $\mu$ V @ TP8/1200 ms	1136/1200 (64 ms)	1064/1200 (136 ms)	4.12b

**Table 4.14:** Information about the effect clusters associated with the three interest predictors for the genitive segment. Summary information is based upon the electrodes at which the effects are maximal. Abbreviations: 'Ant.' = 'anterior'; 'Post.' = 'posterior'; 'Midl.' = 'midline'; 'Pos.' = 'positivity'.



**Figure 4.9:** Nominalisation segment - Top half: model maps for the 's.AI' predictor for selected time windows. Row a: left anterior positivity in time window from 462–562 ms, with polarity inversion peaking at left posterior electrode P5. Row b: model maps for later time window (998–1098 ms), with right central focus of polarity inversion. Bottom half: mean-curve plot for low, mid and high AI levels at electrode F9. Orange rectangles within the plot demarcate the time windows used for plotting the model maps in the top half.

0.61  $\mu V$  at 620 ms. Note that figure 4.10a illustrating this effect shows the model maps based on a time window slightly preceding the time point of the peak effect which coincides with the time period of minimal probability (i.e. highest significance). The spatial discontinuity of the anterior midline and left posterior positivities suggests that they actually represent the foci of two separate effect clusters. This impression is strengthened by inspection of the spatiotemporal evolution of these two clusters as illustrated by the respective animated



**Figure 4.10:** Genitive segment: Model maps for the 'c.Linking' predictor. Row a: Left-posterior positivity peaking at P5. Row b: Anterior midline positivity peaking at AFz. Row c: Late posterior midline positivity peaking at CPz.

model maps.<sup>6</sup>

<sup>6</sup>All model maps referred to here and in the following are available as supplementary materials at <http://doi.org/10.18452/19291>; see the 'readme.txt' files available with the supplementary materials and section 4.2.6 for more information.

Figure 4.10c shows the model maps of a third effect of the 'c.Linking' predictor which temporally follows the previous two and exhibits a clearly distinct distribution, with a broad centroparietal positivity and a polarity inversion over anterior electrodes bilaterally. This posterior midline positivity extends from 800<sub>pw</sub>/ 992<sub>BONDS</sub> ms to the end of the segment and reaches its peaks of 1.36  $\mu$ V at 1060 ms at electrode CPz. The animated scalp maps show how the distribution underlying the first two effects illustrated by figures 4.10a/b shifts to the posterior midline positivity from briefly before 800 ms post onset. This shift suggests that the time around 800 ms coincides with the beginning of a new processing (sub-)phase.

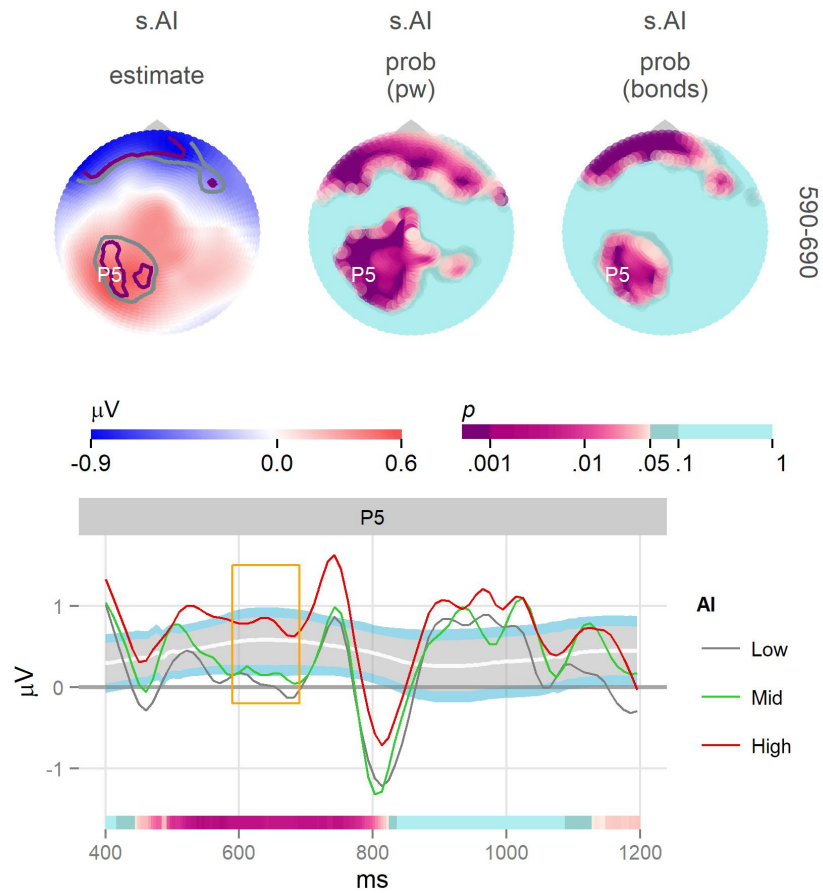
Note that the temporal extent and the timing of the peak effects of the anterior midline and left posterior positivities are consistent with the temporal characteristics of the P600/late positivities. Hence, I will refer to these as 'late' positivities, while I will use 'trailing' in referring to effects falling into the time window from around 800 ms onwards.

Figure 4.11 illustrates the effect associated with the 's.AI' predictor, which shows a centroparietal positivity with a frontal polarity inversion. The posterior positivity has a left-hemispheric focus peaking at electrode P5 at 640 ms and 0.58  $\mu$ V; at P5, it extends from 364<sub>pw</sub>/ 444<sub>BONDS</sub> ms to 860<sub>pw</sub>/ 824<sub>BONDS</sub> ms, which largely coincides with the time range of the late 'c.Linking' positivities. Note, however, that the effect reaches significance again at P5 toward the end of the segment in the time window of the trailing positivities; as the distribution of this effect stays rather stable throughout the whole period, this later effect is not displayed separately. Also, the animated model maps<sup>6</sup> show that according to the pointwise probabilities, this effect is actually consistently flagged as significant throughout the late and trailing time windows at parietal midline electrodes.

Figure 4.12 shows the effects associated with the 'c.Linking:s.AI' interaction predictor for two relevant time windows. The first two effect clusters occur within the time window of the late positivities associated with the 'c.Linking' predictor and are both illustrated by 4.12a, which shows that the distribution underlying the effects in this time window is largely identical to the distribution associated with the late positivities related to the 'c.Linking' predictor, and that the spatial foci of the two effect clusters coincide with those of the 'c.Linking' effects. According to the pointwise probabilities, the interaction is significant at frontal midline electrodes with a maximum at electrode Fz (maximum effect 1.75  $\mu$ V at 548 ms); this large effect is confirmed by the MULTIBONDS, even though this single cluster is interrupted by a short non-significant interval, yielding the two sub-intervals detailed in table 4.14. The second, weaker interaction effect cluster is associated with a left posterior positivity, which ranges from 336 to 824 ms (maximum effect of 0.54  $\mu$ V at 740) ms and just fails to be flagged as significant by the MULTIBONDS.

In addition to these late positivities, another effect cluster arises within the time window of the trailing effects towards the end of the segment. Figure 4.12b illustrates the distribution associated with this trailing effect, which resembles that of the trailing positivity arising with the 'c.Linking' predictor (see figure 4.10c). The focus of the posterior positivity lies at right-hemisphere electrode TP8, where the effect starts at 1064<sub>pw</sub>/ 1136<sub>BONDS</sub> ms and lasts until the end of the segment, where it peaks at 1.36  $\mu$ V.

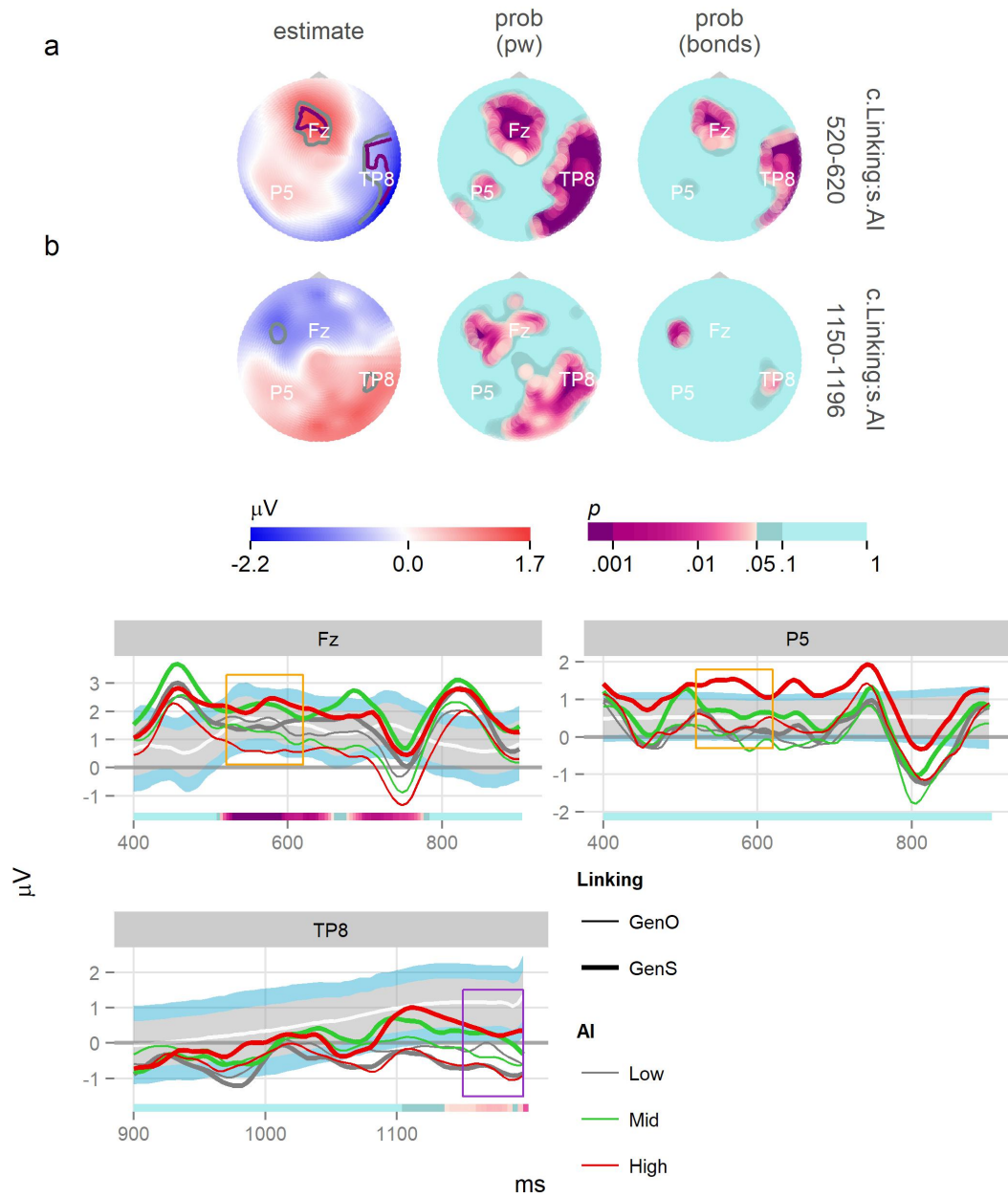
The ERP patterns underlying the interaction effects within the time window of the late pos-



**Figure 4.11:** Genitive segment - Model maps for the 's.AI' predictor: left-posterior positivity peaking at P5. The curve plot at the bottom shows the time window from 400 to 1200 ms post-onset.

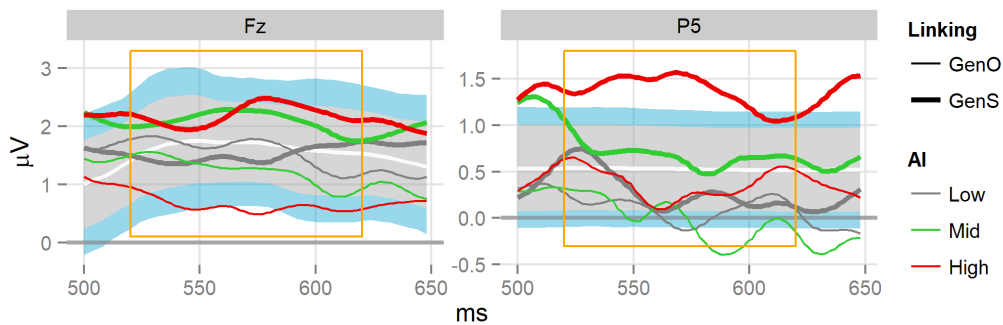
ivities are illustrated by the corresponding curve plots of electrodes Fz and P5 in figure 4.12, and figure 4.13 zooms into the temporal region from 500 to 650 ms. These two figures suggest that there are slightly different ERP patterns at these peak electrodes: the frontal midline patterns at electrode Fz suggest that an increase in the AI correlated with a more *positive* going ERP for GenS linking, while it was associated with a more *negative* going ERP in the GenO condition, with ERPs for low-range AI trials of either condition being about equal within this time period. In contrast, an increase in the AI correlated with a large positive increase of the ERP for GenS continuations at P5, while the pattern for GenO trials seems less clear.

The impact of the AI in each the two 'Linking' conditions is summarised by table 4.15, as determined by the planned follow-up assessments. Figure 4.14 illustrates the effect clusters occurring within the time window of the late positivities: in the GenO condition, the 's.AI' predictor is associated with an effect in this time window which exhibits a scalp distribution



**Figure 4.12:** Genitive segment - model maps for the 'c.Linkings.s.AI' interaction predictor. Row a: anterior midline positivity and left-posterior positivity; maps represent the means of posterior quantities of the time window marked by the orange rectangles in the two topmost curve plots. Row b: model maps for late posterior positivity, based on the time window marked by the purple rectangle in the bottom curve plot.





**Figure 4.13:** Genitive segment: ERP interaction patterns at electrodes Fz and P5. The plot zooms into the time window from 500 to 650 ms.

Predictor	Component	Maximum	Start/End (Length)		Fig.
			BONDS	pointw. <i>p</i>	
AI.GenO	Ant. midl. neg.	-1.22 $\mu$ V @ Fp1/632 ms	404/796 (392 ms)	312/840 (528 ms)	4.14a
AI.GenS	Ant. midl. pos.	1.01 $\mu$ V @ FCz/576 ms	512/700 (188 ms)	492/736 (244 ms)	4.14b
	Left post. pos.	0.84 $\mu$ V @ P5/644 ms	364/856 (492 ms)	260/1200 (940 ms)	4.14c
	Post. midl. pos.	1.14 $\mu$ V @ CP2/1200 ms	1056/1200 (144 ms)	764/1200 (436 ms)	4.15

**Table 4.15:** Information about effect clusters associated with the simple slopes of 's.AI' for the genitive segment. Summary information is based upon the electrodes at which the effects are maximal. Abbreviations: 'Ant.' = 'anterior'; 'Post.' = 'posterior'; 'Midl.' = 'midline'; 'Pos.' = 'positivity'; 'Neg.' = 'Negativity'.

with an anterior negativity and a centroparietal polarity inversion; within this distribution, the negativity is significant at anterior midline electrodes, with the maximum occurring at fronto-polar electrode Fp1; the focus of the positivity lies over right temporoparietal electrodes, as shown in figure 4.14a. The frontal negativity starts to become significant at electrode Fp1 at 312<sub>pw</sub>/ 404<sub>BONDS</sub> ms and lasts until 840<sub>pw</sub>/ 796<sub>BONDS</sub> ms, with the maximum effect of -1.22  $\mu$ V at 632 ms. This frontal negative shift is consistent with the above observation of more negative-going ERPs with rising AI in the GenS condition at the focus of the frontal midline interaction cluster (see figure 4.13).

The second row of table 4.15 provides the details for the frontal midline effect cluster arising in the GenS condition, which shows a somewhat more posterior focus around electrode FCz, where it lasts from 492<sub>pw</sub>/ 512<sub>BONDS</sub> ms to 736<sub>pw</sub>/ 700<sub>BONDS</sub> ms, reaching its peak of 1.01  $\mu$ V at 576 ms. Figure 4.14b illustrates this frontal positive shift and also shows that it is accompanied by a second left posterior positive focus which is maximal at electrode P5, where

it is significant from 260<sub>pw</sub>/ 364<sub>BONDS</sub> ms to 1200<sub>pw</sub>/ 856<sub>BONDS</sub> ms and reaches its maximum effect of 0.84  $\mu$ V at 644 ms. The resulting scalp voltage distribution of the 's.AI' predictor in the time window chosen for figure 4.14b is very similar to the distributions associated with the late positivities of the 'c.Linking' and the 'c.Linking:s.AI' interaction predictor (see figures 4.10a/b and 4.12a). However, the left posterior positivity of 's.AI' in the GenS condition starts as early as 260 ms when assessed with the pointwise probabilities, which precedes the onset of the anterior positivity; figure 4.14c shows the estimates for the time period from 420–460 ms, i.e. briefly before the anterior positivity starts to become significant. The separate spatial foci of the left posterior and the anterior midline positivities and their temporally asynchronous evolution again suggest that the scalp distribution illustrated by 4.14b is the result of the overlap of two late positivities, an impression which is strengthened by inspection of the animated model maps<sup>6</sup>. Importantly, the animations also show that the 's.AI' predictor in the GenO condition is *not* associated with left posterior effects.

Note that according to the pointwise probabilities, the left posterior positivity arising with 's.AI' in the GenS condition around P5 extends until the very end of the segment; in fact, the focus of this late left posterior effect starts to shift to electrodes around the posterior midline, resulting in an effect within the time window of the trailing positivities with a scalp distribution as shown in figure 4.15. While P5 is part of this trailing posterior positivity (at least based upon the pointwise probabilities), this effect shows a more bihemispheric distribution, with a slight right-hemispheric focus around CP2, where it starts at 764<sub>pw</sub>/ 1056<sub>BONDS</sub> ms and lasts until the end of the segment, where it also peaks at 1.14  $\mu$ V.

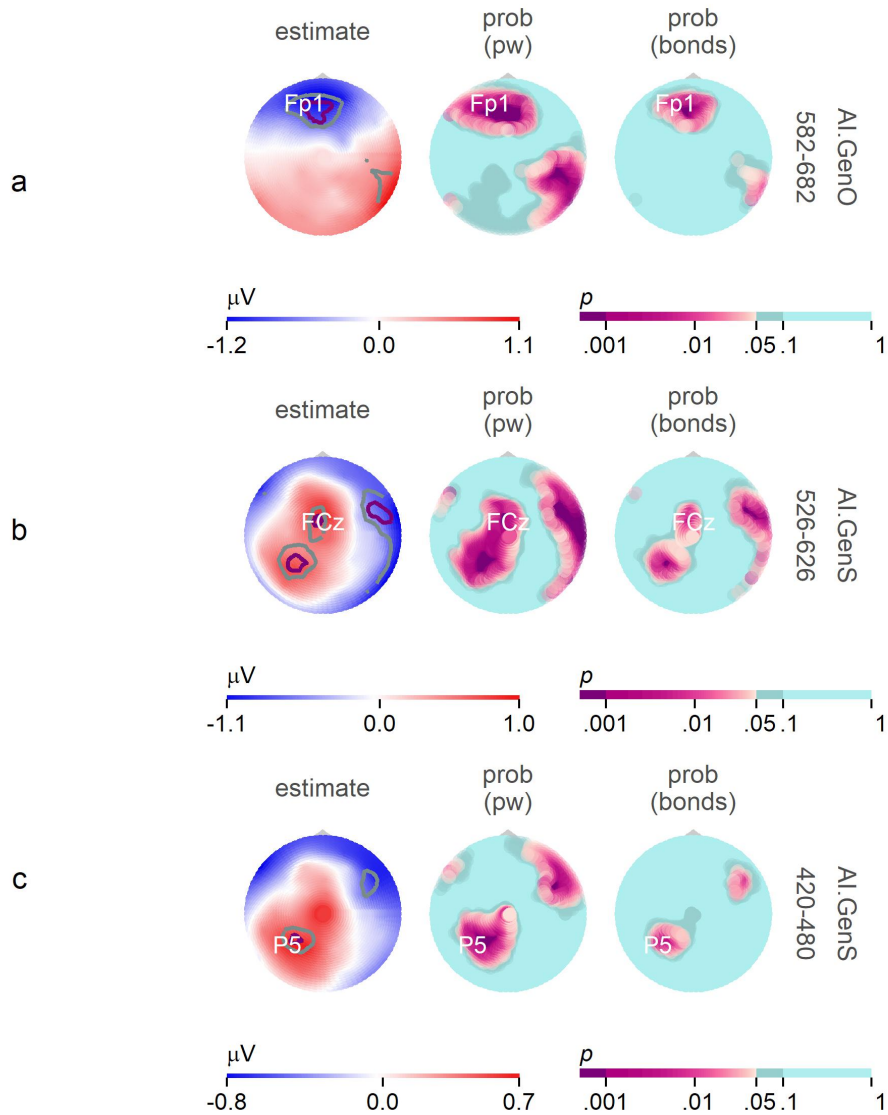
## 4.2.8 Discussion

### Rating data

The results for the rating data are again fully consistent with the findings of the previous two experiments, fully replicating the interaction of the 'Linking' and 'AI' variables and the characteristic interaction pattern. Importantly, this replication of the previously identified acceptability patterns provides the precondition for isolating relevant ERP effects which may partially underlie their generation.

### ERPs in the nominalisation segment

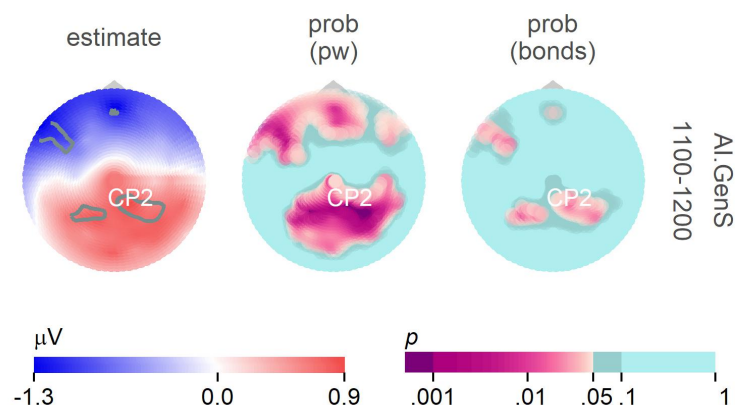
In the nominalisation segment, the 'c.AI' predictor was associated with a left anterior positivity accompanied by a posterior negativity (figure 4.9a/b). An interpretation of this effect is hampered by a number of factors: first, the significant cluster identified as the left anterior positivity is spatially very localised and the global scalp distribution varies somewhat over time, with a shift of clusters within the posterior negativity. The narrow spatial focus of the (not very pronounced) effect and the somewhat diffuse scalp distribution may also be the consequence of the potentially counteracting factors identified in the discussion of the



**Figure 4.14:** Genitive segment - follow-up model maps for simple slopes of 's.AI'. Row a: anterior midline negativity for 'AI' in the 'GenO' condition, peaking at Fp1. Row b: anterior midline positivity for 'AI' in the 'GenS' condition, peaking at FCz. Row c: Left posterior positivity for 'AI' in the 'GenS' condition, peaking at P5; note that the time window for this model map slightly precedes the time of the peak effect at P5.

SPR experiments in section 4.1.7, including prior priming of the nominalisations (or their base verbs) by the context sentence, imageability and durativity.

Second, the most severe difficulty is posed by the lack of a parallel behavioural effect in the SPR experiment, which may allow to interpret and characterise the effect on a sound



**Figure 4.15:** Genitive segment: Follow-up model maps for simple slope of 's.AI' in the 'GenS' condition in the very late time window. The effect maximum occurs at electrode CP2.

basis. Without such a solid behavioural footing – and in the absence of relevant ERP findings from previous studies which clearly relate to the thematic focus of the current experiment – it is actually not possible to determine the direction of the observed effect and thus the polarity of the shift. Without a behavioural basis showing that an increase in the 'AI' variable correlates either with an increase or decrease in RTs on the nominalisation, the anterior effect could represent a positive or a negative shift. In principle, an anterior effect may in fact be in line with the findings of Steinhauer et al. (2001) (see also Malaia et al., 2009) and the predictions of Brennan and Pytkänen (2010). However, Steinhauer et al. already faced a similar challenge and pointed out the difficulty of soundly interpreting the anterior ERP effect associated with lexical-semantic complexity in their study in the absence of behavioural evidence. These authors also discussed the possibility that the anterior negativity for count nouns they observed may as well be viewed as a positivity associated with mass nouns. As a consequence of these issues, I will refrain from further interpreting the effect witnessed.

On a final note, I would like to point out that the model only tested for linear effects of 's.AI' and other predictors included; an inclusion of appropriate non-linear terms may have revealed more pronounced and/or additional effects. However, it would also have complicated the interpretation of such effects, especially since an inspection of the RTs obtained in the SPR experiment did not provide any support for possible non-linear patterns on the behavioural level (see figure 4.5 above).

## ERPs in the genitive segment

### N400

The WFMM-based analysis did not reveal a relevant modulation of the N400, thus failing to confirm the hypothesised role of the N400 for the resolution of the linking conflicts involved in the present paradigm, as stated by hypothesis 1. While some of the effects isolated

started within the core time window usually associated with the N400 (300–500 ms; see table 4.14), they extended beyond this time interval and reached their peaks clearly later than 400 ms post onset. As discussed in the following, the observed effects are rather consistent with interpreting them as late positivities.

Haupt et al. (2008) discuss the possibility that the lack of an N400 effect in some relevant studies may be the consequence of component overlap. Thus, some experimental paradigms are prone to eliciting task-related P300 effects, including those with explicit tasks to be executed after sentence presentation; in such cases, the P300 will overlap with the N400 and may (partially) cancel out N400-related effects. While this may of course be a possibility with the present experiment, which involved an explicit task, this question cannot be resolved here. Note that the remarks of the previous section about potentially non-linear effects and the failure of the model specified to identify these apply equally here; again, the interpretation of possible non-linear effects may have been non-trivial, since the hypotheses formulated do not refer to non-linearities and the relevant behavioural RT patterns did not exhibit pronounced non-linearities.

The discussion of possible implications of the absence of an N400 modulation will be postponed until after the discussion of the effects revealed in the later time windows. I will come back to this issue in the subsequent joint discussion.

### **Late positivities**

The WFMM-based analysis of the genitive segment revealed a number of effect clusters whose characteristics are consistent with classifying them as late positivities; these include the anterior midline positivities observed for the 'c.Linking' effect and the 'c.Linking:s.AI' interaction and the respective left posterior positivities, which peaked at electrode P5 and also comprised an effect of the 's.AI' predictor (see table 4.14). These start around the peak or end of the N400 time interval and reach their maxima in a time window from 548–740 ms, which corresponds well with the temporal profile of standard P600 effects. Note that I use the term 'late positivities' here without deeper theoretical implications (such as the eADM's distinction between late positivities and P600 effects proper – see section 3.2.4.2), but simply as referring to positivities occurring within the standard P600 time window. As I will argue in the following, the interpretation of these effects as late positivities also receives support from their spatial distributions and the specific associated ERP patterns, which are consistent with the behavioural RT data obtained in the SPR experiment.

Importantly, I have claimed that the effects occurring within this time window (see figures 4.10a/b and 4.12a) constitute two distinct late positivities, rather than one extended cluster. This claim is based upon the independent evolution of the two clusters' spatial foci around anterior midline and left posterior electrodes and the specific, distinct patterns underlying the 'c.Linking:s.AI' interaction.

Starting the discussion of these two late positivities with the anterior midline cluster, it is worth noting that the anterior spatial focus of this effect deviates from the centroparietal distribution usually associated with 'standard' P600/late positivity effects. Nevertheless,

positivities with an anterior focus or involvement within this late time window have previously been found in sentence processing experiments (e.g. Carreiras, Salillas, & Barber, 2004; Friederici, Hahne, & Saddy, 2002; Hagoort & Brown, 2000; Kaan & Swaab, 2003; Wang, Bastiaansen, Yang, & Hagoort, 2012), and while several authors have proposed that the different scalp distributions of anterior and posterior late positivities warrant distinct functional interpretations, there has been no agreement on how to best characterise either. In their survey of ERP markers related to syntactic processing, Hagoort, Brown, and Osterhout (1999), for example, argue that the posterior positivity indexes processes underlying the parsing of syntactically incorrect sentences, whereas the anterior positivity is related to discarding the preferred structural analysis of a sentence and reassigning a less preferred one. This interpretation is in principle compatible with the findings of Friederici et al. (2002), who obtained a (post-N400) posterior positivity for syntactic violations and a positivity with an anterior focus for syntactically complex sentences. Such a division of labour between the two types of positivities, however, is not fully consistent with the findings of Kaan et al. (2000); in this study, the processing of structurally more complex sentences resulted in a posterior positivity. In the experiment of Kaan and Swaab (2003), ungrammatical and dispreferred structures elicited a posterior positivity, while the complexity manipulation resulted in modulation of the anterior positivity. Given the nature of their specific sentence material, Kaan and Swaab raise the possibility that the anterior positivity may index processing difficulty at the discourse level. This interpretation, in turn, is contested by Burkhardt (2006), who rather associates discourse integration processes with a (left) posterior positivity (see also section 3.2.6).

The question of the possible role of such late anterior positivities in sentence processing has received renewed attention based upon the results of a number of recent studies, even though from a somewhat different perspective (e.g. DeLong, Quante, & Kutas, 2014; DeLong, Urbach, Groppe, & Kutas, 2011; Federmeier, Wlotko, De Ochoa-Dewald, & Kutas, 2007; Thornhill & Van Petten, 2012). Most of these studies were concerned with the predictability of given words within a sentential or discourse context, a topic which has generally been the domain of research on the N400. Yet, recent studies and reviews have shifted attention on different types of late positivities sometimes occurring as part of a biphasic N400-late positivity pattern in such experiments, in which such positivities have materialised with either a posterior, or an anterior distribution. Van Petten and Luka (2012) provide a systematic review of such 'post-N400 positivities' (PNPs) which suggests an interesting relationship between the nature of the critical predictability manipulation and the manifestation of a (potential) PNP with a posterior or anterior distribution. Van Petten and Luka concentrated on comparing the type of PNP arising with critical words which either constituted unexpected and incongruent sentence completions (i.e. conceptual anomalies), or highly unexpected, but conceptually congruent completions. In the latter case, expectedness was quantified via cloze probabilities of the respective critical words, as determined by offline sentence completion studies.

While a PNP occurred only with about one third of all analysed studies with incongruous sentence completions, the topography of these PNPs was parietal in the majority of cases

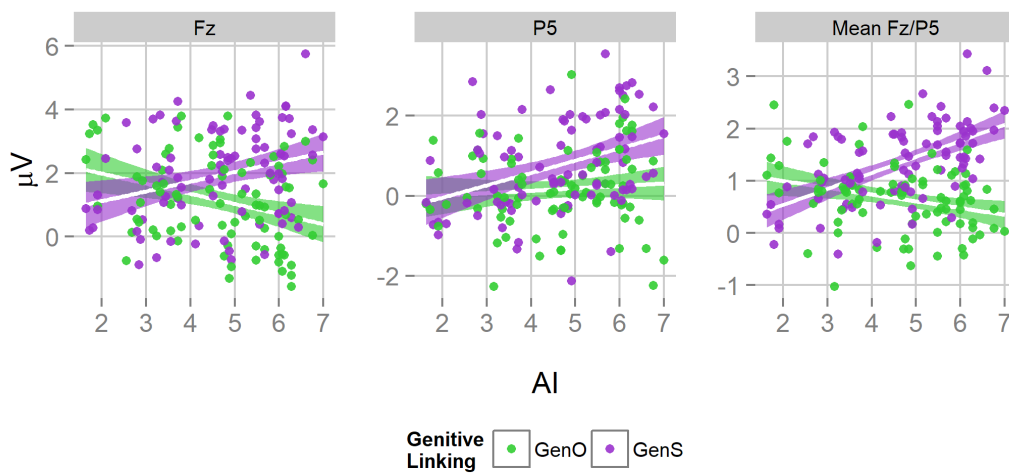
(17/21). When the critical word was highly unexpected, but conceptually congruous, though, a PNP appeared quite frequently and exhibited predominantly an anterior scalp topography. Discussing the different contexts and factors which seem to correlate with modulations of the N400, posterior and anterior PNPs, the authors suggest that these components index differential functional processes: on their view, the amplitude of the N400 reflects the *benefits* reaped from supportive conceptual-semantic context in integrating a given word, which corresponds well with the standard interpretation of the N400. The posterior PNP can be seen as an instantiation of the same monitoring- or reanalysis-related P600s associated with syntactic processing difficulties or 'semantic illusions' (see section 3.2.3 for a brief review of N400 and P600-related issues), which accounts for the preponderance of posterior PNPs with conceptually incongruous sentences. The anterior PNP, however, indexes the *cost of a failed prediction*, where Van Petten and Luka understand the term 'prediction' as implying that the parser is expecting a *specific* lexical item; when such a prediction is disconfirmed, an anterior PNP frequently ensues. The exact eliciting conditions are still not fully understood, but crucial factors seem to include a strongly constraining sentence context with a high cloze probability for a specific word – which allows to form such a specific prediction in the first place – and the conceptual congruity of the alternatively presented, unpredicted word with the preceding sentence context (for discussion, see DeLong et al., 2014; Thornhill & Van Petten, 2012; Van Petten & Luka, 2012).

Speculating about the cognitive functions indexed by such anterior PNPs, Kutas (1993) offered an early interpretation in terms of *inhibition* of predicted words which were subsequently not presented, a suggestion discussed by Van Petten and Luka (2012) and DeLong et al. (2014). DeLong et al. (p. 160) note:

“One tentative explanation for the type of processing indexed by the frontal positivity may have to do with inhibition. For instance, Levy and Anderson (2002) in a review of inhibitory processes and semantic memory retrieval, argue that executive control mechanisms (in prefrontal cortex) may be required to override prepotent responses when a memory must be selectively retrieved in the face of other competing memories. [...] At a very minimum, the rough correspondence (taken at face value) of the gross prefrontal brain regions implicated in inhibition networks with the scalp regions where we observe the frontal positivity is tempting.”

This inhibition-related view of late anterior positivities fits well with the suppression mechanism outlined above in the discussion of the SPR results as well as with the specific pattern of results observed in the present experiment. In effect, the pattern found is exactly that predicted for the N400 by hypothesis 1, even though with reversed directionality: as the follow-up analysis of the simple slopes of the 's.AI' predictor revealed, the interaction effect observed at anterior midline electrodes indicated a significant positive shift of the ERP with increasing AI in the GenS condition and a significant negative shift in the GenO condition (see figures 4.13 and 4.14 and table 4.15). The plot rendered in the leftmost panel of figure 4.16 shows linear smooths for the 'AI' variable fitted separately at the GenO and GenS

levels, based upon mean ERP values computed within the time window around the peak of the anterior interaction effect, from 520–620 ms. Note that the pattern closely resembles the RTs observed in the 'PostGenitive 1' region in the SPR experiment (see figure 4.4): the slopes of the estimates are roughly symmetric and the ERPs for low-AI items are about equal in the GenO and GenS conditions (with slightly more positive ERPs in the GenO condition, which is again in line with the RT results). Thus, the patterns underlying the anterior interaction positivity also meet the second part of hypothesis 1, which predicted roughly equal ERPs at the low end of the AI-range, and are further fully consistent with the RT patterns witnessed in the SPR experiment.



**Figure 4.16:** Genitive segment: Linear smooths for 'AI' at each level of 'Linking', computed on basis of mean ERP values within each nominalisation separately for 'GenO' and 'GenS' continuations. Mean values were calculated in the time window from 520-620 ms post-onset at electrodes Fz and P5; the last panel shows the smooths based upon the mean values across these two electrodes.

Importantly, a functional interpretation of late anterior positivities as indexing inhibitory processes is immediately compatible with the concept of the ESR, which I put forward as a tentative component of a toy competition model regulating the linking of GenO and GenS with German *-ung*-nominalisations in section 4.1.7. From such a perspective, the anterior positivity identified can be considered functionally close to anterior PNPs and can be interpreted in the present context as reflecting additional inhibitory resources which are mobilised in an attempt to suppress the (implicit) co-argument of the genitive encountered, as outlined above. Note that the temporal extents of the anterior effects summarised in table 4.14 are consistent with previous reports, which suggest that such anterior positivities may start relatively early, within the N400 time window, and frequently extend until approximately 900–1000 ms post onset (see DeLong et al., 2014; Thornhill & Van Petten, 2012, for example). Thus, the spatial and temporal characteristics of the observed anterior effects as well as a tentative functional interpretation as an index of inhibitory processing are fully consistent with



recent work on anterior PNPs.

In section 3.2.6 I pointed out that the current experiments also involve the processing of discourse-related aspects and discussed the Nref as a component which has been interpreted as a processing correlate of referential ambiguity at the discourse-level (Van Berkum, 2012; Van Berkum et al., 2007). While the spatial and temporal characteristics of the Nref are similar to those of the anterior positivity found in the current experiment, the polarity is reversed. Thus, if the anterior effect observed indexed the resolution of a – rather broadly construed – discourse-related referential problem reflected in an anterior Nref, the direction of the effect should be reversed: the more difficult the integration, the more negative the ERP should become. The observed patterns thus do not support such an interpretation.

Another discourse-related ERP marker discussed in the previous chapter is the late posterior positivity found by Burkhardt (2006), which the author views as reflecting processes of re-organising or updating the reader's established discourse model. Interestingly, this posterior positivity's scalp distribution showed a left posterior asymmetry with a focus around electrodes CP5/P3, which is in line with the topography of the left posterior positivity identified in the present study (peaking at P5). Note that the specific ERP pattern underlying the left posterior interaction effect is consistent with such a discourse-related interpretation: the ERPs displayed in the right panel of figure 4.13 show a clear linear positive shift with increasing AI in the relevant time window for GenS continuations, but a rather flat and less distinct pattern for GenO continuations; in addition, the ERP arising with low-AI GenS continuations was of roughly equal amplitude as (high-AI) GenO items. While the left posterior interaction effect cluster was less pronounced than the anterior one and was only flagged as significant by the pointwise probabilities (table 4.14), the follow-up analysis of the simple 's.AI' slopes clearly showed the differences between the anterior and left posterior effects, with the latter being restricted to a positive shift in the GenS condition and no relevant effect with GenO continuations (see table 4.15 and figure 4.14). This pattern is summarised by the plot in the panel in the middle of figure 4.16, with the linear smooths of the mean ERPs at P5 within the 520–620 ms time window showing a more positive-going ERP with increasing AI for GenS continuations and a flat slope in the GenO condition.

Thus, the left posterior ERP pattern observed is coherent with hypothesis 2 as well as Burkhardt's functional interpretation of this late positivity: on this view, it indexes processes related to checking the coherence of the encountered predicate-argument combination with the previously established discourse representation. The pattern observed corresponds to the coherence component of the toy model outlined above (see figure 4.6c) and may indicate that GenOs and low-AI GenS continuations are perceived as equally coherent with the discourse representation. Coherence declines linearly with an increasing AI for GenS continuations, however, resulting in complete incoherence with maximal AI-values.

While a causal relationship between the two late positivities found and the degree of acceptability observed cannot be established on basis of the current data, it is nevertheless tempting to speculate that the eventual acceptability may be a composite quantity derived mainly from these two ERP indexes. In the outline of the toy model, I raised the possibility that the degree of acceptability may be approximated as the mean of the ESR and coherence

components, as displayed in figure 4.6d. Following the above suggestions that the anterior and left posterior ERP effects may functionally correspond to these two components, respectively, the rightmost panel of figure 4.16 shows the smooths for the mean of the ERPs at electrodes Fz and P5 displayed in the other two panels. The resulting pattern indeed strongly resembles the acceptability data obtained in all of the three experiments conducted.

### **Trailing positivities**

The two late positivities were followed by a positivity with a posterior midline topography, which differed markedly from the distributions of the preceding positivities and exhibited spatial foci around centroparietal electrodes, emerged at around 800 ms post onset and lasted until the end of the segment (table 4.14 and figure 4.10c). During this time interval, which I referred to as 'trailing', GenS continuations resulted in a more positive ERP than their GenO counterparts; towards the end of the segment, this effect was partially accompanied by an interaction of the 'c.Linking' and 's.AI' predictors (see figure 4.12b); the follow-up analysis again showed that the 's.AI' predictor was associated with a significant posterior positive shift (table 4.15 and figure 4.15).

Note that the interpretation of these effects is hampered by the fact that they temporally overlap with the processing of the subsequent word, which was presented 600 ms after the genitive argument. Nevertheless, the possibility that late positivities/P600 effects may reflect the processes of two temporally adjacent phases has been discussed previously, for example by Hagoort et al. (1999), who note that "it has been reported that for syntactic violations the initial phase (500-750 ms) of the P600/SPS has a fairly equal scalp distribution, whereas the second phase (750-1000 ms) shows a clear parietal maximum [...]. It thus could be that the P600/SPS is not one effect but a family of effects, with the (additional) contribution of different generators in the early and late phase of the effect, related to functionally different states of the parser" (p. 293) (also see Kaan et al., 2000, for example). If the functional interpretations of the late positivities presented above are on the right track, then the trailing positivities may be understood as reflecting some kind of rather general monitoring and/or rechecking process, which is consistent with more recent functional descriptions of late positivities, as discussed in section 3.2.3 (Van Herten et al., 2005). Such a view of the trailing positivity effects is consistent with the specific patterns observed and may also relate to the eADM's Generalised Mapping step (section 3.2.4.2), which is also assumed to monitor for ill-formed or marked constructions. The additional processing cost associated with such a check may also contribute towards the degree of acceptability attributed to a given structure.

### **Joint discussion**

The most relevant results of the present ERP experiment are the absence of a modulation of the N400 and the presence of a number of later positivities in the genitive ERP segment. With respect to the N400, one possible scenario I pointed out above is that the null result may

be the consequence of concurrent positivities which may have cancelled out relevant N400 modulations (see Haupt et al., 2008). Beyond this possibility, which cannot be assessed further here, the reason for this absence of N400 effects may be found in one or more of the differences between the current paradigm and those which the eADM's view of the N400 is mainly based upon, the most relevant of which I discussed in section 3.2.4.3.

Within the eADM, the N400 is thought to act as the primary marker of the Compute Prominence and Compute Linking processes, the two central steps related to predicate-argument linking in the eADM (Bornkessel-Schlesewsky & Schlesewsky, 2009; Haupt et al., 2008). As outlined previously, the two processes are housed in separate processing streams within the eADM's architecture: while both take place in the time window associated with the N400 (phase 2 of the eADM), Compute Prominence is part of the processing branch entered with non-predicating arguments, while Compute Linking is triggered within the predicate branch. Typically, the processing correlates of either have been measured in such a way that the category of the critical words was consistent with the domain of the respective stream. The present experiment, however, involved accessing and exploiting semantic properties of argument NPs originating within the predicate stream, thus cross-cutting the two processing streams. The absence of N400 effects may thus imply that information which is rooted in the opposing stream can indeed not be accessed during phase 2; the modulations of the late positivities, on the other hand, may be understood as being in line with the idea that information rooted in the opposing stream may only be acted upon during phase 3.

Alternatively, the results may also imply that the semantic information which has to be accessed in the resolution of the linking conflict at stake in the present context is not within the domain of the processes driving relevant N400 effects. On this view, the lack of N400 modulations may not be due to a separation of two processing streams within the N400 time window; rather, the principled differences between predicate-induced prominence properties on the one hand and NP-inherent prominence cues encoded in natural prominence scales on the other hand (see sections 1.3.2 and 3.2.4.3) may call for different coping strategies for linking conflicts stemming from either type of prominence information. Note that it is not immediately clear why such substantial differences should exist, given that both types of argument prominence information can be understood as resulting in (semantically) more or less distinct co-arguments, as stated by the eADM's principle of Distinctness. While this principle has mainly been formulated with respect to NP-inherent prominence features taken from natural prominence scales, it may similarly apply to verb-induced prominence information. After all, the current interpretation of the anterior late positivity in terms of inhibitory processes may in principle also be valid in the context of conflicts driven by NP-inherent prominence features, where suppression of 'intruding' co-arguments may equally require inhibitory resources. Why an anterior late positivity should occur in one case, but not the other (for example as part of a biphasic N400-late anterior positivity pattern), is thus a non-trivial question. Thus, the current findings certainly underscore the assertion of Bornkessel-Schlesewsky and Schlesewsky (2009) that more research on potentially relevant prominence scales and the way they interact in sentence processing is needed. Note that I will return briefly to this issue in the general discussion presented in the following chapter.

Other potentially relevant factors I pointed out previously include the nature of the task, which did not involve the establishment of a linking configuration in the first place, but rather checking whether a given construction is consistent with an established target linking. This also entailed accessing non-local information represented on the discourse level. Either of these factors may have contributed to the current results, including the lack of an N400 effect. Ultimately, this issues arose as a consequence of some of the peculiarities of the current paradigm involving nominal linking, which certainly raise questions about the generalisability to other contexts, especially verbal argument linking processes; note that such wider implications are discussed within the concluding chapter 5.

Apart from the absence of N400 effects, the most noticeable positive finding of the ERP experiment are the patterns observed with the late positivities, which I have argued to reflect different functional processes: the evaluation of discourse coherence for the left posterior positivity and inhibitory activity for the anterior positivity. Together with the trailing positivity witnessed, the current results fit well with the above mentioned observation that late positivity/P600 effects may show an internal biphasic organisation, with an initial broad positivity followed by a positivity with a parietal focus, indexing different neural and functional operations across these earlier and later phases (e.g. Hagoort et al., 1999; Kaan et al., 2000). The current claim that the late positivity effects found reflect two functionally distinct processes is of a somewhat different kind, relating to the spatiotemporal and functional fine structure of the earlier of these two phases.

Generally, it is assumed that such late effects as the N400 or P600 reflect the joint activity of an ensemble of multiple active neural sources. While it would go far beyond the scope of this thesis to attempt a discussion of the internal organisation of late positivities, I would like to point out that the picture which has emerged is a coherent one: on the one hand, the specific and distinct ERP patterns associated with the anterior and left posterior foci are internally consistent by their correspondence with key components of the toy model sketched on basis of the behavioural RT results; on the other hand, they are crucially also consistent with recent interpretations of these ERP markers as reflecting inhibitory activity and discourse-related processes, respectively, as established in independent research (Burkhardt, 2006; DeLong et al., 2014; Thornhill & Van Petten, 2012).

The present interpretation of the late positivity effects as mapping onto two functionally distinct and concurrent spatiotemporal effect clusters (the anterior midline and left posterior clusters) also raises the question of the degree of information exchange between these. At least if the competition model sketched above is taken as basis, the final assessment of discourse-level coherence crucially depends upon the output of the upstream components, including the ESR, which I have tentatively linked to the anterior positivity. Since the anterior and left posterior effects develop largely in parallel, an assumption to be made is that the neural sources driving the left posterior effect must have received information about the outcome of the competition process (the dominance component) within the temporal window associated with the late positivities. This is another conjecture which cannot be substantiated here, and the picture is of course further complicated by the subsequent trailing positivity, which may further have added to the eventual outcome.

Focussing on the anterior positivity, the patterns observed may also be viewed from a related, but slightly different perspective, on which the anterior effect is a reaction to the degree to which a particular genitive was *predicted* on basis of the context sentence and the preceding nominalisation. Such a view would imply a somewhat different interpretation of the positivity which equates its eliciting conditions with those of the studies investigating the impact of plausibility and prediction on the N400/PNP complex reviewed above (DeLong et al., 2014; Thornhill & Van Petten, 2012; Van Petten & Luka, 2012, for example): thus, the anterior positivity may still reflect inhibitory processes, but they would be conceived of as a kind of a more general, reflex-like response, rather than a more specific reaction of the parser, here conceived of as mobilisation of the ESR.

While it might be argued that within such a scenario the degree of expectedness of either GenO or GenS of a specific co-argument pair may be derived purely on the basis of the frequency of the co-occurrence of either argument with the respective nominalisation, I doubt that such a purely frequency-based view would stand an empirical test. The modulation of the late anterior positivity by the AI is graded and roughly linear in the GenO as well as the GenS condition and a frequency-based explanation of these patterns appears unlikely: though a low AI may well correlate with higher frequency of GenS realisation in a language corpus, a graded decline of GenS occurrence with increasing AI along the scale seems improbable. Rather, I would expect a sharp drop in GenS frequencies in mid-AI regions already, with frequencies close to or at zero. Further, it is rather unlikely that GenS realisations are roughly equal in number to GenO continuations for minimum-AI nominalisations. The relatively low frequencies of GenS realisations in natural speech may relate to the fact that the most natural way of expressing a transitive verb's subject is within a PP headed by *durch* (*by*), which unambiguously identifies the subject as such, as pointed out in section 2.2. Of course this reflects my personal intuitions and the only way to resolve this issue would be to conduct a comprehensive corpus study and enter the resulting frequencies into the statistical model.

Nevertheless, even if a purely frequency-based explanation of the ultimate source of the prediction-related view of the (inhibitory) anterior positivity may not prove successful, this alternative perspective is still compatible with a semantics-driven approach to deriving the degree of expectedness of either genitive argument. From this angle, the anterior positivity still reflects a kind of reflex to the degree to which a given genitive was predicted given the prior context (rather than allocation of ESR), but the degree of anticipation stems from the semantic properties of the arguments, as determined by the lexical-semantic properties of the base verb in form of the AI variable. The basis for the parser's predictions would thus still be provided by the affectedness-induced prominence of GenO (or, alternatively, its semantic prototypicality and the semantic distance between the co-arguments), as described above, but the functional view of the anterior positivity would have a somewhat different quality than that viewing it as the electrophysiological correlate of the hypothesised ESR component.

Interestingly, support for the tentative interpretation of the anterior positivity as reflecting (something like) ESR mobilisation, which I put forward above, can be found in ERP studies outside of the area of language processing. Folstein and Van Petten (2011) carried out an

ERP study on the categorisation of visual stimuli in which participants had to classify cartoon creatures as belonging to one of three categories labelled 'Mogs', 'Nibs' and 'Others'. These creatures were composed of three different task-relevant features comprising global body shape, colour and body markings, and the Mogs and Nibs categories both had a central prototype (e.g. a red striped horse or a purple spotted fish). During the experiment, participants had to assign presented stimuli to one of the three categories, instructed to follow a 'two-out-of-three' rule. The critical comparison was between items belonging to the 'far' and 'near' boundary conditions: stimuli belonging to the far boundary condition drew all three relevant features from their own category, thus acting as prototypical items; stimuli of the near boundary condition, on the other hand, only had two features of their own category, while the third one was compatible with a different class, thus falling near the boundary between Mogs and Nibs.

Crucially, the comparison of items of the far and near boundary conditions showed that the latter elicited a more pronounced late positivity at anterior electrode sites (see also Azizian, Freitas, Parvaz, & Squires, 2006; Folstein & Van Petten, 2004, for compatible results with similar paradigms). Folstein and Van Petten (2011, p. 839) link this anterior positivity to the engagement of executive processes and speculate that the "[c]oordination between working- and long-term memory may be the core process driving frontal positivities". These results thus show intriguing parallels to the present findings, especially when the current paradigm is viewed from the perspective of argument prototypicality and the semantic distance of the co-arguments, as discussed previously. Indeed, on this view the tasks involved in these studies share core aspects with the present paradigm, which may include the need to actively suppress representations which 'intrude' on the currently processed item as a consequence of an increase in distance in the relevant feature space between a stimulus and its prototype, which often goes hand in hand with an increase in similarity to other categories. While it remains to be worked out whether Folstein and Van Petten's speculative interpretation of the anterior positivity and the present one viewing it mainly as an index of inhibitory activity may be compatible, the findings just reviewed may provide external support for my interpretation of the present results as an electrophysiological correlate of the ESR. If on the right track, the anterior positivity observed may be understood as a general marker of higher-level inhibition processes of a domain-general nature; these may arise as a consequence of suppressing representations which compete with a currently processed item due to proximity in a relevant feature space with central prototypes.

Of course, such an interpretation of the anterior positivity remains highly speculative for the time being and raises further questions, such as its relationship to other ERP components and its neural substrates. DeLong et al. (2014), for example, call attention to the parallels between the anterior and posterior positivities they observed and the P3a and P3b components, respectively, two varieties of the P300. They refer to the framework of Polich (2007), who views these effects as indexing inhibitory neural activity. Whether or how the present anterior positivity is related to the P3a and whether parts of Polich's proposal are compatible with the inhibition-interpretation of the anterior positivity remains to be investigated.

Several authors note that the frontal spatial focus of the late anterior positivity is roughly

consistent with neural sources in prefrontal brain regions implied in executive functions, including inhibitory processes (see DeLong et al., 2014; Folstein & Van Petten, 2011, for example). Levy and Anderson (2002) review aspects of inhibitory control processes, highlighting the roles of the prefrontal cortex (PFC) and anterior cingulate cortex (ACC), and Thompson-Schill et al. (2005) focus on the left ventrolateral prefrontal cortex (VLPFC) as a substrate for general regulatory processes which also figure prominently in different aspects of language processing. As pointed out in section 3.2.5, the findings of Hindy et al.'s (2012) fMRI study on dissimilarity-based interference induced via object state-change (see example (115)) as well as the results of the MEG experiment on similarity-based interference in sentence processing of Leiken and Pykkänen (2014) are consistent with an involvement of left inferior frontal brain regions in the resolution of such conflicts. Whether the late anterior positivity observed in the studies reviewed above and in the present experiment can indeed be traced back to sources within (left) PFC – and/or other regions implied in the executive control network such as the ACC – again remains an open question. Further investigations involving the reconstruction of ERP sources may shed some further light on this issue.

In the discussion of the late anterior positivity witnessed in the ERP study on highly expected words of Federmeier et al. (2007), the authors rightly point out that it is not obvious whether the anterior positivity in such experiments functionally equals those anterior positivities previously observed in sentence processing experiments involving some sort of syntactic integration difficulty. However, Hagoort et al. (1999, p. 293) put forward a tentative interpretation of anterior positivities arising with the latter type of processing conflicts as reflecting “the processing costs associated with overwriting the preferred or most activated structure”. Thus, if anterior positivities are indeed related to the suppression of initially preferred syntactic analyses, a common functional view of such anterior effects as indexing inhibitory activity may be plausible across the different experimental contexts.

As a final point, I will return to the lack of clear effects of the AI variable in the nominalisation segment, which brings up questions about the basis of the ERP patterns witnessed with the genitive segment – after all, I have argued that the latter ultimately depend upon the lexical-semantic properties of a nominalisation's base verb. The absence of (interpretable) effects in the nominalisation segment is consistent with the null-result in the respective regions in the SPR experiment, which may have been a consequence of weaknesses in the experimental design and confounding factors (e.g. imageability and durativity) in both cases. Brennan and Pykkänen (2010) also failed to obtain effects of lexical-semantic complexity in an MEG experiment (section 3.2.1); however, considering that they found clear effects of semantic complexity in a previous behavioural study using the same stimulus material, an explanation in terms of confounding factors is less likely. Brennan and Pykkänen speculate that the null-result may have been due to the limited sensitivity of MEG to specifically oriented cortical sources or their choice of time windows and ROIs for the analysis of the MEG data. In the best case, future studies may prove that such methodological shortcomings are valid explanations for the present elusiveness of verbal lexical-semantic complexity. If these remain elusive, however, a more complicated scenario may open up, calling for alternative explanations of the current effects observed with the resolution of the linking conflict in the genitive

segment. Since these involve issues of a more general nature, I will briefly come back to possible implications in the concluding chapter.

### 4.3 Summary

I have presented two online sentence reading experiments on the impact of the degree of affectedness on the processing of lexical-semantic aspects of deverbal predicates and on the linking of genitive arguments. The rating data of the SPR and the ERP experiment were fully consistent with the data of the acceptability judgement study summarised in chapter 2 and reproduced the pattern underlying the interaction of the 'AI' and 'Linking' variables. With respect to research question one formulated in the Introduction, the analysis of the RT data obtained in the SPR experiment failed to confirm the hypotheses about the effect of lexical-semantic complexity for the processing of the nominalised predicate; similarly, the analysis of the ERP data of the nominalisation segment did not yield strong support for the hypothesised impact of affectedness on the processing of the underlying verbal semantic structure. In both cases, methodological shortcomings may have contributed towards this null-result. With respect to research question two, the RT and ERP data associated with the genitive segment, on the other hand, showed a clear interaction of 'AI' and 'Linking'; the interaction pattern was largely consistent with the acceptability data and in part confirmed the hypotheses about the role of affectedness during the online resolution of the argument linking conflict, which were partially derived on basis of the eADM (Bornkessel & Schlesewsky, 2006; Bornkessel-Schlesewsky & Schlesewsky, 2009). The most noticeable results of the ERP experiment were a lack of N400 modulations and the presence of distinct interaction effects for two concurrent late positivities, one with an anterior midline focus and one with a left posterior spatial maximum. While the absence of N400 modulations clashed with the hypothesis about the role of the N400 for the resolution of prominence-based conflicts derived from the eADM, I argued that it may still be compatible with the model, given the eADM's architecture and the specifics of the present paradigm. The pattern of the late left posterior positivity cluster is consistent with viewing it as a reflection of discourse-level coherence evaluation. Interestingly, the pattern underlying the interaction of the anterior positivity corresponded to the predictions about the N400, and I tentatively suggested that it may reflect inhibitory resources mobilised in the attempt to regulate the present linking conflict, an interpretation which is consistent with the ESR component of a toy model of the competition mechanism sketched on basis of the RT patterns as well as recent research on the effect of prediction in language processing (DeLong et al., 2014; Thornhill & Van Petten, 2012; Van Petten & Luka, 2012, for example). This interpretation is further consistent with research on prototype-based classification of non-linguistic stimuli (e.g. Folstein & Van Petten, 2011); the findings within this line of work also show promising intersections with the present thesis by providing support for the prototype-based view of the linking conflict investigated here. From this perspective, the anterior positivity may index inhibition of competing representations which intrude upon the item to be processed due to proximity in a feature space in which central prototypes are defined, possibly of a domain-general kind. This interpretation and the frontal spatial fo-



cus of the positivity is further compatible with prefrontal neural sources implied in executive processes. For the time being, however, these interpretations remain speculative and raise further questions, some of which are addressed in the following chapter.



## Chapter 5

# General discussion

The three main questions this thesis started off with as listed in the Introduction are repeated here (with research question one referring to effects on the nominalisation segment and research question two to processes arising with the genitive segment):

1. Does affectedness have any consequences for the processing of predicates with verbal semantic structure? If so, can we isolate behavioural and electrophysiological processing correlates of this semantic property on the predicate itself?
2. Does predicate-induced affectedness have any impact on the integration and licensing of NP arguments following the predicate within a sentence? If so, can we identify behavioural and electrophysiological processing correlates of such processes on the respective argument NP?
3. How do the findings fit with current models of sentence processing which address these aspects from related perspectives?

In the following, I will return to these questions and address them given the experimental results in the previous chapters; I will discuss a number of issues and possible ways of addressing some of these in future studies and, where possible, I will put them into a somewhat broader context.

### 5.1 Generalisability

Before I address these three initial research questions, there is one critical point to be clarified and discussed: in the present work, I have concentrated on exploring the above questions using linking phenomena in NPs, i.e. within the nominal domain, but I have not provided any evidence that the results may indeed generalise to other linguistic contexts. The choice of using deverbal eventive nominalisations in the present work was based upon a number of crucial characteristics of processes regulating argument linking in German *-ung*-nominalisations,

such as the irrelevance of case or agreement marking. The key consideration was that these should act together to minimise the influence of other, morphosyntactic cues, maximise the need to make use of semantic information and thus also maximise potential processing effects related to accessing this information and operating on it (see section 2.6). The results and their present interpretation are in line with the theoretical view of the relationship between verbal semantic structure and its reflexes in derived eventive *-ung*-nominalisations as well as with the impact of affectedness on nominal argument linking in the constructions used (see chapter 2). Given this consistency, I have implicitly assumed that the processing correlates identified should be involved in, and generalise to, verbal argument linking processes too. The present studies, however, do not prove this, since they did not test for parallel effects in the verbal sentence domain.

Thus, a crucial aim for possible future research following up on these issues should be the replication of the findings in the verbal domain, i.e. in full sentences with a verb-argument complex, to show that the findings generalise from the current niche context to a more general setting. After all, the differences between verbal and nominal argument linking processes (e.g. in German the relevance of morphosyntactic features on NP arguments and verbs, such as case and agreement marking) may have consequences for the role affectedness ultimately plays for the integration of arguments during sentence processing: for example, it is conceivable that the absence of any other cues in the experiments presented here – which I hypothesised to maximise potential effects – may actually have rendered the impact of affectedness disproportionately big. The parser may have treated it as a kind of ‘last resort’ information source to resolve the arising conflict; it is conceivable that this information source is not tapped to the same degree – or not at all – in presence of explicit morphosyntactic cues.

These crucial issues can be resolved by subsequent experiments which investigate argument-related reanalysis processes in dependence of the verb-induced affectedness level using more common types of paradigms and experimental materials. The examples in (129) illustrate very simple, syntactically transitive sentences which involve local grammatical function reanalysis: the first argument NP *die Frau* (*the woman*) is case-ambiguous between a nominative subject and an accusative object, but due to its sentence-initial position it is initially interpreted as the subject; this interpretation must be revised when the sentence-final NP *der Mann* (*the man*) bearing nominative case is encountered.

- (129) a. Die                                      Frau    bewunderte der                                      Mann.  
           the.FEM.SG.NOM/ACC woman admired    the.MASC.SG.NOM man  
           ‘The man admired the woman.’  
       b. Die                                      Frau    tötete der                                      Mann.  
           the.FEM.SG.NOM/ACC woman killed    the.MASC.SG.NOM man  
           ‘The man killed the woman.’

This is true independently of the verb *bewundern/töten* (*admire/kill*), but these verbs of course induce very different degrees of affectedness on the object. A prediction derived from the findings of the experiments presented here would be that the affectedness degree

modulates processing effects related to the reanalysis, even if in more subtle ways than in the present experiments. Note that effects during sentence processing related to semantic properties of the predicate may in general be of a rather subliminal nature, such as the processing difference between eventive and stative verbs found by Gennari and Poeppel (2003) and illustrated by example (103) or the effect of telicity on the reanalysis of reduced relative clauses as in (104) (see Malaia et al., 2011; O'Bryan, 2003). While the garden-path effect arising due to the reanalysis is strong in both sentences in (104), the telicity effect is much more subtle.

- |       |  |               |
|-------|--|---------------|
| (103) | a. The retired musician built his second house from scratch. | <i>event</i>  |
|       | b. The retired musician loved his second child very much.    | <i>state</i>  |
|       |  |               |
| (104) | a. The actress spotted by the writer left in a hurry.        | <i>telic</i>  |
|       | b. The actress chaperoned by the writer left in a hurry.     | <i>atelic</i> |

One external key principle for deriving hypotheses in cases such as (129) is the eADM's principle of Distinctness discussed in section 3.2.4.1, which states that the parsing system assumes that "[t]he participants in an event should be as distinct as possible from one another in terms of all available dimensions of prominence" (Bornkessel-Schlesewsky & Schlesewsky, 2009, p. 44). When it comes to the role of prominence features of NP-arguments, remember that the eADM's focus lies on NP-inherent characteristics such as animacy and definiteness. Provided that Distinctness also has consequences for predicate-induced semantic properties of argument NPs in such contexts and given my interpretation of the present findings (see section 4.2.8), a decrease in prototypicality of the object argument and thus also in semantic distance between the two co-arguments should in these cases result in longer RTs in sentences such as (129a) than in (129b), since the reanalysis process would be made more difficult by the less distinctive semantic profiles of the co-arguments. Within the eADM, the most relevant electrophysiological marker of the Compute Prominence step is hypothesised to be the N400, but the findings of the ERP study presented in this thesis suggest that the degree of affectedness should modulate the late anterior positivity, rather than the N400 or late posterior positivities (which may also occur in such constructions, but should not correlate with the affectedness variable in the same way). If so, the current view of the role of affectedness in online argument linking would be considerably strengthened. Such a replication across the verbal and nominal domain would in fact provide very strong support for the present view of the late anterior positivity as a reflection of inhibitory activity indexing the suppression of co-arguments in dependence of their degree of predicate-induced prototypicality and modulation of their semantic distance.

As the preceding discussion of the ERP findings (section 4.2.8) suggested, 'the' P600/late positivity is not a monolithic phenomenon, but rather may occur with different spatial distributions (with posterior, anterior or global foci) and different temporal subphases, with the first phase (extending until about 800 ms) sometimes showing a broadly distributed positivity, which then shifts to a posterior focus in the second phase (see Hagoort et al., 1999, for

example). Further, the relationship between late positivities and attributed functional interpretations may not be one-to-one, but rather one-to-many (e.g. Bornkessel & Schlesewsky, 2006; Haupt et al., 2008, see also section 3.2.4.2). These observations fit with the results of the present ERP experiment, with different spatial foci (anterior and left posterior) of the late positivities interpreted as reflecting functionally different processes (inhibition vs. assessment of overall discourse-related coherence, respectively). These findings and the present interpretations show that a closer examination of possible sub-processes underlying (different variants of) late positivities should be undertaken, and confirmatory evidence on the present interpretation of the functional significance of the late anterior positivity in the resolution of argument linking conflicts could provide valuable information in this respect.

## 5.2 Research question one

Returning to the main aims of this thesis, question 1 could in principle be answered succinctly with a simple “no”, since no effects of the AI variable on the nominalisation-related segments were found for the RTs and the effects in the ERP data were too feeble and diffuse to justify deeper interpretation. However, this assessment is qualified by a number of weaknesses of the experimental design, as already pointed out previously (see sections 4.1.7 and 4.2.8), including failure to control for a number of potentially confounding variables (e.g. imageability and durativity).

One additional possible factor I have not discussed so far is the heterogeneity of the real world conceptual-semantic fields the nominalised predicates represented, i.e. the type of affectedness scale associated with a given verb (in the sense of Beavers, 2011). In the experiments presented, these included location scales, scales referring to different physical and mental processes and others (see section 2.7.2.1). As discussed in section 1.5.1, in Beavers’ scalar analysis of affectedness, the structure of an affectedness scale – which represents the degree of affectedness as an aspect of grammatically relevant semantic *structure* – is separated from any conceptual aspects carried by a verb (i.e. information belonging to semantic *content*). The main advantage of severing the structure of an affectedness scale from its real-world properties is a high level of parsimony by treating all different types of change in a uniform way.

If such a view of the structure and organisation of affectedness scales is on the right track, it is still not obvious how it may be implemented at the neurocognitive level. An assumption I have implicitly followed is that the scale structure may be handled by a common neural generator or network across different scale types and may thus be measurable in the experimental paradigm applied. However, the relationship between the internal and theoretical organisation of scales and possible neural implementations thereof may be more complex, and mixing scales from different conceptual-semantic fields within an experiment may result in interference with effects related to the degree of affectedness. As suggested by the results of the fMRI study of Kemmerer et al. (2008), among others, aspects of the semantic content of a verb are processed by different neural areas (see section 3.2.1). In ERP studies,

the use of predicates representing different types of scale may thus result in activation of different generators, which may render the identification of effects related to affectedness more difficult. Where possible, subsequent investigations of this issue should thus make use of a single scale type only or account for the type of scale in the experimental design and analysis. A more principled treatment of this factor should be easier when verbs, rather than nominalisations, are used as experimental items, due to more leeway in the construction of experimental items.

Finally, if correlates of affectedness (and related phenomena) should prove elusive in subsequent experimental investigations, such a consistent lack of effects may give support to those who are sceptical about the theoretical necessity and 'psychological reality' of the internal semantic structure of words (de Almeida, 2004; de Almeida & Dwivedi, 2008; Manouilidou & de Almeida, 2013, for example). Under such a scenario, the interpretation of the effects of affectedness underlying argument integration processes found on the genitive arguments (question 2) would have to be reconsidered. de Almeida (2004) argues that processing effects generally interpreted as tapping into the internal semantic structure of lexical items (such as in coercion constructions like *The boy began the book*; section 3.2.1) actually reflect processes of pragmatic inferencing. Such a view may provide the basis for alternative explications of the present results, though I will not further pursue such alternative routes here.

### 5.3 Research question two

The absence of effects related to affectedness on the nominalised predicates may also have been the result of the operationalisation of affectedness as the continuous AI variable, which may not have been appropriate to detect possible patterns. However, the effects of the same AI predictor on the genitive argument segments in the experiments were clear and generally fitted well with the hypotheses about the impact of affectedness on the acceptability of GenO and GenS genitives and processing correlates of their integration.

The most conspicuous ERP correlates of the regulation of the genitive argument NPs' integration were the two late positivities with left posterior and anterior midline foci, which I tentatively interpreted as reflecting coherence assessment at the discourse level and inhibitory activity involved in the regulation of the competition of the two co-arguments, respectively. From the perspective of the original research question about the existence and nature of processing correlates underlying the integration and licensing of NP arguments in dependence of predicate-induced affectedness profiles, the most intriguing finding of the present work is probably the anterior positivity, which I already discussed extensively in section 4.2.8. While the functional interpretation I attributed to the anterior positivity in the present context is highly speculative, it nevertheless provides a strong hypothesis which can be tested empirically.

Thus, if experiments testing for the generalisability of the current findings to other contexts provide further evidence for the involvement of the anterior positivity in the regulation of ar-

gument integration/reanalysis processes related to semantic features of arguments induced by verbs, subsequent studies may shed more light on its relation to similar positivities arising in other contexts. Two of these I already discussed are the anterior PNPs arising with congruent but highly unexpected lexical items in sentences (see DeLong et al., 2014; DeLong et al., 2011; Federmeier et al., 2007; Thornhill & Van Petten, 2012; Van Petten & Luka, 2012) and the anterior positivity found in experiments on prototype-based classification of visual stimuli (Azizian et al., 2006; Folstein & Van Petten, 2004, 2011). If the present results stand the test of reproduction, follow-up experiments could directly compare anterior positivities elicited by these two paradigms with anterior positivities involved in argument linking processes in the same group of subjects. Parallels between these positivities across these different experimental contexts may strengthen my interpretation of the present anterior positivity's functional significance in terms of inhibitory processes involved in the suppression of representations competing with a processed item due to proximity in a feature space with prototypes.

Remember that invoking connections to prototype-based approaches to argument linking (e.g. Ackerman & Moore, 2001; Dowty, 1991) is not an ad hoc move, but rather a natural view which is opened up by Beavers' scalar approach to the role of affectedness in argument linking and its explicit links to Prototype theories, as worked out in Beavers (2010) and summarised in section 1.5.1. Evidence for common neural markers and substrates involved in prototype-based classification or integration processes in linguistic and non-linguistic (Azizian et al., 2006; Folstein & Van Petten, 2004, 2011) paradigms may thus also give experimental support for argument linking approaches along the line of Beavers (2010, 2011), Dowty (1991), Ackerman and Moore (2001) and others.

In the discussion of the initial acceptability judgement experiment in section 2.7.3.7, I already pointed out a possible way of further investigating the role of prototype-features of NP genitive arguments for the nominal linking patterns investigated in the present thesis: this involves a comparison of the current unipolar view, which focuses on the degree of affectedness of the argument realised as object only, and a bipolar perspective which incorporates prototype properties of the subject in addition (e.g. Blume, 2000; Grimm, 2011). The latter approach would enable the computation of a kind of 'semantic distance index' as a potentially more comprehensive indicator of the degree of semantic transitivity of a verbal predicate; such a variable could not only yield valuable additional information with the present experimental paradigm, but also with possible follow-up experiments probing for generalisability of the current findings to other linguistic contexts, and may provide supportive evidence for the prototype-based view of the role of predicate-induced semantic factors in argument linking.

Further issues arising with the interpretation of the anterior positivity observed include its possible connections to other potentially related processing markers and the influence of the task administered. Among the former is the AMF, a processing correlate which has been found in a number of MEG studies on sentence processing and interpreted as a marker of operations underlying semantic composition (such as coercion, see Pykkänen et al., 2011; Pykkänen, Martin, et al., 2009; Pykkänen & McElree, 2007; Pykkänen, Oliveri, & Smart, 2009, for example), as outlined in section 3.2.3. The anterior midline distribution of the AMF



and its possible role in the processing of semantic aspects of sentences invites speculation about common functional operations and neural sources with the anterior positivity found in the present study. While comparisons may reveal potential parallels, these are rendered less likely by the fact that studies using EEG have failed to reproduce the anterior effects found in the MEG experiments of Pykkänen and colleagues (see Baggio et al., 2010; Kuperberg et al., 2010). Kuperberg (2007) discusses the possible role of the specific task administered to subjects in evoking more or less prominent late positivities with conceptually incongruent sentences, focussing on the possibility that tasks involving explicit acceptability judgements are more likely to result in such effects (see also Kolk et al., 2003; Schacht, Sommer, Shmuelovich, Martíenz, & Martín-Loeches, 2014). Since the present experiments involved acceptability judgements of a rather complex kind, it may well be that a more passive kind of task significantly modulates the anterior positivity.

## 5.4 Research question three

The final research question posed at the beginning of this thesis concerns possible implications of the experimental findings for models of sentence processing. Taking another look at research question three, you may notice that the question could have been formulated more generally, by extending its scope to 'current models of sentence processing', without any further restrictive clauses. However, at the present moment, I believe that such a general and overarching assessment is too early, given that proof about reproducibility and generalisability to processes involved in argument integration around verbs – i.e. in the sentential domain – is still to be provided, as pointed out above.

Nevertheless, I have made use of one particular recent model of sentence processing for building a context for predictions and discussion of the experimental results, the eADM (e.g. Bornkessel & Schlesewsky, 2006; Bornkessel-Schlesewsky & Schlesewsky, 2009). As discussed in more detail in section 3.2.4, the present work has a lot in common with the eADM in terms of the aims and the general perspective taken, including the interest in semantic determinants of argument linking processes, the links to prototype-based approaches to argument realisation, the role of the prominence of NP arguments and the semantic similarity of co-arguments, as expressed in its principle of Distinctness. At the same time, however, there are also crucial differences concerning the type of NP-related semantic information focussed upon (see section 3.2.4.3): while the eADM largely deals with verb-independent and NP-inherent features, such as animacy or definiteness, the present focus lies upon the role of affectedness, a semantic feature of NP-arguments induced by the respective predicate. In discussing the ERP results in the previous chapter, I brought up the possibility that the lack of an N400 effect and the presence of the two late positivities (especially the anterior one) on the genitive argument segment may be due to two factors: first, in the context of the eADM, one explanation may be that when encountering the critical argument NP, the semantic information stemming from the opposing predicate stream cannot be accessed and acted upon during phase 2, but only during phase 3, thus resulting in late positivities, rather than N400

effects. Remember though that Czipionka (2014) found interactions between NP-inherent animacy information and verb-class (prototypically transitive nominative-accusative vs. non-prototypically transitive nominative-dative), which she argues may reflect an interplay of inherent and 'derived' semantic argument properties. Her results indicate that such interactions may also take place earlier on. Thus, despite some important differences in perspective and paradigm to the current work (see section 3.2.2), this question may thus merit more investigation.

Alternatively, I addressed the possibility that the relevant verb-induced semantic information lies without the scope of the processes driving the respective N400 effects in contexts with NP-inherent argument linking conflicts. On this view, while both types of linking conflicts may be regulated by prominence information, it is the principled differences between predicate-induced and NP-inherent prominence information which may account for the patterns observed, though I left open the question of the source of these differences on a neurocognitive level.

Speculating somewhat further about this point, the following extended quote from Levin and Rappaport Hovav (2005, pp. 173–174) addresses interesting differences in the way that semantic attributes of NP arguments figure into the ranking of thematic roles with verb-induced properties on the one hand, and NP-inherent features derived from natural prominence scales on the other hand (such as the animacy hierarchy – also see sections 1.3.2 and 3.2.4.3; emphasis is mine):

“There are two drawbacks in using the personal hierarchy as a basis for ranking semantic roles. First, properties such as person and definiteness (or, perhaps, assumed familiarity), which are often taken to be components of the extended animacy hierarchy, are properties of the fillers of roles and are not event-based properties – that is, they are not derived from entailments which a verb imposes on its arguments by virtue of the parts they play in the event it describes. This can be true of animacy as well, if a verb does not entail the animacy of any of its arguments. These properties, then, should not be relevant to a hierarchy of event-based roles. Second, even when animacy is taken to be event-based, it imposes a rather coarse-grained ranking, since NPs bearing the agent, experiencer, benefactive, and recipient roles, for instance, are all typically animate. Animacy, then, does not impose an exhaustive ranking on all the semantic roles associated with a particular verb, let alone on all roles across verbs. *This observation reflects a more general point: the thematic hierarchy is unlike most natural prominence scales because it ranks arguments on a fundamentally different basis from them. Other scales rank arguments according to the values of a 'simple' attribute, such as animacy or definiteness – an attribute whose values fall along a single dimension. Although a semantic role may be considered a value of an attribute of arguments, it is not a 'simple' attribute. At best, a semantic role can be viewed as defined by a cluster of properties, such as Dowty's proto-role entailments*

[...], with each property representing a value of a simple attribute. [...] *No single attribute imposes a ranking on all semantic roles; rather, each role can only be characterised in terms of the value of several attributes.*"

This perspective opens up an intriguing parallel to the context eliciting anterior positivities observed in the ERP studies on the classification of visual stimuli which more or less match their own prototype or fall near the boundary of other categories (Folstein & Van Petten, 2004, 2011). Folstein and Van Petten (2011, p. 834) point out that their predictions about frontal ERP effects for stimuli with features of distinct categories were based upon "the hypothesis that categorization problems that cannot be solved after detection of single features or simple conjunctions trigger a secondary analysis that is more demanding of executive functions dependent on prefrontal cortex"; further, they note that their categorisation task shared some crucial characteristics with previous ERP studies on episodic memory which also yielded similar late prefrontal positivities, stating that "single stimulus attributes do not signal the correct response, which must be based on attribute conjunctions" (Folstein & Van Petten, 2011, p. 839).

Accordingly, one factor triggering the anterior positivity found in the present ERP experiment may be rooted in the differences between NP-inherent semantic features based upon natural prominence scales and event-based, verb-induced NP-properties worked out in the above quote of Levin and Rappaport Hovav: while rankings based upon the former involve 'simple' attributes, hierarchies of thematic (Proto-)roles derived from the latter are based upon (clusters of) several attributes. In parallel to the conjecture of Folstein and Van Petten, the present anterior positivity may thus be the result of the need to mobilise prefrontal areas to resolve the particular linking conflict. While the only variable investigated here was the AI of the argument realised as object, an assumption which could be made is that linking processes involving verb-induced semantic prominence features of NP arguments nevertheless always involve a comparison of *all* relevant attributes which jointly characterise the respective (Proto-)role, thus triggering the anterior positivities observed. On this view, linking conflicts based upon NP-inherent attributes like animacy, as typically investigated in the eADM, on the other hand, may be resolved without such an extensive recruitment of prefrontal sources, resulting in N400 effects, for example.

The present findings certainly underscore the call of Bornkessel-Schlesewsky and Schlesewsky (2009) for further studies investigating how many and which different types of prominence scales figure into argument linking operations in sentence processing, and if and how these interact. Future investigations may thus combine NP-inherent and predicate-induced semantic argument prominence attributes to yield data addressing these issues in detail. The present work is also consistent with the general spirit of the eADM by stressing the necessity of implementing theoretical insights about how verbal semantic structure and semantic characteristics of arguments are organised and about the way they feature in the regulation of argument linking during language processing, in particular the role of prototype-features as well as semantic transitivity, which closely correlates with affectedness, and its various conceptions (see section 1.5.2.1). Deeper insights into these issues will also contribute to the understanding of the various ERP markers involved.

With respect to general implications for models of sentence processing, I believe that the main message to take away from the present work – even without proof of reproducibility and generalisability available at the moment – is that the issues raised and investigated in this thesis are worth pursuing in the first place. With more information about the role of predicate-induced semantic prominence information on NP arguments in online argument linking at hand, further questions certainly relate to the way possible findings may complement or challenge prevalent views of the part of the syntax-semantics interface which deals with verbs and their arguments in sentence processing. These are often based upon the traditional notion of thematic roles as atomic entities with concrete labels such as Agent and Patient (cf. Tanenhaus et al., 1989), a theoretical approach which suffers a number of shortcomings, as discussed in section 1.3. While I fall short of addressing these larger topics here, I hope that I may have provided some inspiration for further investigating them.

## 5.5 Other issues

Finally, there are a number of interesting issues concerning possible implications of the present findings for theoretical accounts of argument linking in the nominal domain, especially with German *-ung*-nominalisations. Since resolving these is not among the aims of this thesis, I will briefly address them here without going into further detail. One point I already brought up in the discussion of the initial acceptability judgement study (see section 2.7.3.7) relates to possible implementations of the graded nature of the effects observed in the present study, which were present in all three experiments conducted. I suggested that a way to implement the graded acceptability patterns may be to translate the implicational Affectedness Hierarchy of Beavers (2010, 2011) into a ranking of constraints as done in the framework of OT (Prince & Smolensky, [1993]/2004), similar to the way in which semantic properties of verbs have been shown to capture effects on auxiliary selection (see Legendre, 2007; Sorace, 2000; Sorace & Keller, 2005).

While such a mapping could account for graded acceptability patterns, it would still constitute a gradedness based upon discrete levels of a hierarchy derived from linguistic reasoning – in the present work, however, the most relevant patterns found are based on the AI computed from intuitive judgements of naive speakers, yielding a numeric, continuous and more fuzzy measure. An interesting issue arising is thus the relationship between such continuous and fuzzy variables and those based upon theoretical constructs with discrete levels. In the present context, legitimate questions to be asked include similarities and differences in the ways the continuous AI and a discrete multilevel-variable reflecting Beavers' original implementation of degrees of affectedness within the Affectedness Hierarchy are able to capture acceptability patterns and processing correlates. Going still somewhat further, one may also ask how graded acceptability and processing effects based upon such continuous, fuzzy variables derived from ratings of semantic aspects of linguistic units are reconcilable with the usually discrete nature of constructs used in linguistic theory. Two examples of other studies which successfully used the former type of variable to reveal processing effects are

Hindy et al. (2012) and Coll-Florit and Gennari (2011, see section 4.2.8): in both cases, continuous predictors based upon subjective ratings of semantic properties of verbs or verb-argument complexes (implied change of state and temporal duration of events described by verbs, respectively) revealed relevant processing correlates. While I will not further go into this topic, note that the following part II of this thesis introduces two analytical software packages which can be used to address such questions from an empirical point of view, by providing easy access to statistical methods which can model the effects of such variables for different kinds of psychophysiological curve data, including ERPs.

Another point I addressed in the discussion of the results of the acceptability judgement study is the fact that with an increase in the AI, acceptability decreased strongly with GenS continuations and increased significantly, albeit less pronounced, with GenO continuations, a pattern replicated in the acceptability data of the two processing experiments presented subsequently. The latter finding contrasted with my initial prediction that acceptability should be about equally high with GenO continuations throughout the range of the AI, without a significant effect of the AI predictor in this condition, based upon the work of Ehrich and Rapp (2000). However, I argued that the significant improvement in acceptability with an increase in the AI for GenOs can be explained by assuming a crucial role of semantic prominence of the GenO argument in terms of its degree of affectedness – or related concepts such as proximity to the Proto-Patient role – in a competitive suppression mechanism: on this account, GenOs can always be linked as postnominal genitive arguments, but their eventual acceptability is still somewhat influenced by the respective AI, due to a decrease in the ability of GenOs to completely suppress competing GenS co-arguments with lower AI values.

This conjecture is in line with the results of the two processing experiments, given the toy model of the competition mechanism sketched in section 4.1.7 and the functional interpretation of the two late positivities within this model (section 4.2.8). Resolving the respective interaction patterns showed that these two processing markers differed mainly in the significance of the slope of the AI predictor in the GenO condition: while it correlated with a significant negativisation with GenOs for the anterior positivity, it did not significantly modulate the ERP with the left posterior cluster. I tentatively argued that the patterns underlying the left posterior late positivity are compatible with understanding it to reflect processes of checking the coherence of the respective predicate-argument complex with the discourse representation established in first context sentence, corresponding to the coherence component of the toy model; the anterior positivity was interpreted as indexing mobilisation of inhibitory resources, representing the toy model's ESR component, and the eventual acceptability ratings can be approximated by the mean of these two measures. In principle, the different patterns underlying these two late positivities may thus be understood to reflect rule-driven operations – in case of the left posterior positivity – vs. 'compensatory' processing operations, in case of the anterior positivity. If such a differentiation is on the right track, it may have implications for empirical attempts to isolate the rules at work with argument linking in *-ung*-nominalisations and to capture them from a theoretical point of view, since judgements may in fact index a complex composite measure reflecting the influence of different rule-based and processing factors.

## Conclusion

In the first part of this thesis, I have investigated the question how predicate-induced semantic prominence of NP arguments impacts online argument linking processes, including the processing of the predicates themselves and the integration of arguments following the predicate. I have focussed on the role of the degree of affectedness of a syntactically transitive verb's direct object, making use of deverbal, eventive *-ung*-nominalisations and genitive arguments. Effects of affectedness for the processing of the deverbal predicates themselves (research question one) have remained elusive, which may have been due to a number of methodological issues with the experimental paradigm applied. The results for the integration of the genitive arguments (research question two), on the other hand, showed the interaction effects as largely predicted on basis of the semantics-based approach to argument linking with *-ung*-nominalisations of Ehrich and Rapp (2000) and the graded approach to affectedness worked out by Beavers (2010, 2011). The relevant interaction pattern was found in all experiments and surfaced in acceptability judgements, RTs and ERPs. The most conspicuous finding was the modulation of a late anterior positivity in the ERP data, which I tentatively interpreted functionally as reflecting inhibitory activity to suppress the intruding co-argument. On this view, predicate-induced semantic prominence of NP arguments correlates with proximity of an argument in a feature space with central prototypes such as Proto-Agent and Proto-Patient, and the present findings may be explained by assuming a competitive linking mechanism, as sketched in the toy model presented. This account is in line with prototype-based approaches to argument realisation (see Beavers, 2010; Dowty, 1991, for example) and semantic transitivity (e.g. Blume, 2000; Grimm, 2011). Further, the present interpretation of the late anterior positivity offers promising links to recent studies on this ERP component within the area of sentence processing (e.g. DeLong et al., 2014; Thornhill & Van Petten, 2012; Van Petten & Luka, 2012) as well as the classification of visual stimuli in the context of central prototypes (Folstein & Van Petten, 2004, 2011). I have discussed possible implications of these findings for the eADM, a model of sentence processing which takes a related perspective, and I have pointed out a number of issues arising with the interpretation of the present results, foremost among which is the crucial question of reproducibility and generalisability to other linguistic contexts. With hindsight, this thesis has probably raised more questions than it may have answered – if I have been successful in at least bringing up these questions for potential future work, I think it has been worth writing this thesis and, hopefully, also reading it. The following second part now introduces the two R packages which have built the basis for analysing the ERP data presented here.

## **Part II**

### **Got curves?**

**Introducing two R-packages for  
WFMM-based analysis of curve-  
and ERP-data**





## Chapter 6

# Backgrounds: towards holistic analyses of curve data

The current chapter opens the second, methodological part of this thesis, which introduces two R packages interfacing with the wavelet-based functional mixed model (WFMM) developed by Morris and Carroll (2006). The first of these packages, *wrapfmm*, is a user-friendly R wrapper for the core WFMM software and facilitates WFMM-based analysis of general curve data considerably. The *stepmom* package (spatiotemporal electrophysiological model maps) builds upon *wrapfmm* and offers a suite of functions for exploring multi-electrode ERP-data and analysing them using WFMMs without the need to specify temporal windows or spatial regions of interest ('ROIs'). While detailed information about these packages will be provided in the following two chapters, here I will put them into context, discussing some methodological and conceptual backgrounds of current standard approaches to curve- and ERP-data analysis and of some more recent alternatives.

As pointed out in the preface, at the time of publication of the present thesis, the package versions as presented in the following are not in sync with current R versions and thus actually outdated for most purposes. Thus, the practical value of this second part and the package versions presented mainly lies in introducing the general approach to the analysis of curve data used in more detail, and in providing the background for interested readers to be able to more comprehensively understand the ERP analyses applied in chapter 4.2.6 in the first part (including, for example, the computation and interpretation of the simultaneous posterior probabilities used for inference on effects, as presented in section 7.7.2) and some discussion of related issues. However, the packages introduced will soon be complemented by much more comprehensive versions that will offer many more features; for more details on these aspects and availability of the packages as presented in the following, please refer to the preface. Since these packages will maintain many of the principles and functions presented here in the respective chapters in similar form, the two package-specific chapters 7 and 8 can also be considered as complementary background readings for these.

The chapter begins with a brief review of the rise of standard linear mixed effects models (LMMs) for the analysis of scalar data in (psycho-)linguistics and of their limits when applied to complex curve- and event-related potentials (ERP) data. Following a survey of a number of recent alternative approaches to ERP data analysis, I first discuss functional data analysis (FDA) as a recent framework for analysing general curve data and then outline the WFMM as a particular FDA tool, discussing some of its key aspects and features in detail. Throughout these sections, I comment on the potential of LMMs and (W)FMMs as 'holistic' modelling tools, in the sense that they contribute towards analysing scalar and functional data by making use of the full amount of structure present in the data at various levels, thus minimising the loss of information for estimating any parameters of interest. While doing so, I discuss some practical limitations and challenges for holistic modelling approaches for curve data in the context of the WFMM.

## **6.1 Limitations of the traditional ANOVA approach to the analysis of scalar data**

### **6.1.1 The challenge: the 'language-as-a-fixed-effect-fallacy'**

Experimental linguistic studies involve testing human participants with language materials (items), whether these be phonemes, syllables, morphemes, words, phrases, sentences or other units. In both cases, the subjects and items involved in an experiment are (more or less) *random samples* from larger *populations* of subjects and items, respectively. Usually, we are interested in drawing conclusions from the data at hand which are likely to generalise beyond the specific speakers who took part in the experiment. During data analysis, this generalisation of effects beyond the actual subject pool is taken care of by treating the experimental participants as a *random factor*, thus explicitly acknowledging subjects as a random source of variability and accounting for it. On the same grounds, the experimenter wants to make sure that her conclusions do not only hold for the language material applied in the study, but are likely to generalise to a new sample of items and doing so implies to treat experimental items as a random factor too. Thus, only if both of these common random sources of variability are properly accounted for in the statistical analysis conducted, the study's findings can be taken to generalize beyond its immediate set of participants and language items.

The need to include items as a random factor in the analysis of linguistic experiments along with experimental participants was first pointed out by Coleman (1964), though it took a few more years to make an impact in the research community; Clark (1973) took up Coleman's arguments and again made the case for the inclusion of by-subject and by-item variance components in analysis of variance (ANOVA) analyses. Failing to account for random item variability was to commit what Clark termed the 'language-as-a-fixed-effect-fallacy': such an implicit treatment of items as a fixed rather than a random factor would entail the loss of generalisability of a study's results to the 'item population'. Ignoring the random nature of

item variability amounts to confounding the expected mean squares of the treatment effect with random item variability in the numerator when  $F$ -ratios in a by-subject ANOVA are computed. Thus, an  $F$ -ratio indicating a significant effect may imply an actual treatment effect, large item variability, or a combination of these – as a consequence, the Type I error rate (i.e. rejecting  $H_0$  when it is actually true) may exceed the nominal  $\alpha$ -level (Clark, 1973).

### 6.1.2 The ANOVA fix: $F'/F'_{min}$ and the $F_1 \times F_2$ criterion

Since no appropriate error term to test the respective treatment effect against exists for the computation of conventional  $F$ -ratios in ANOVA analyses in which both subjects AND items are considered to be random effects, the technical solution advocated by Clark (1973) was to compute a quasi- $F$  ratios, or  $F'$ , instead. Under some circumstances (such as when faced by missing data, e.g. due to discarded error responses), however, it may become difficult to calculate a quasi- $F$  ratio in practice. In such situations, Clark recommended computing a  $F'_{min}$  value as a lower bound estimate of the quasi- $F$  ratio, which can be obtained by first conducting *separate* subject and item analyses, known as  $F_1$  and  $F_2$ , respectively. For a subject analysis, the data points for each design cell are calculated by averaging over items, while in an item analysis they are computed by averaging over subjects. In a subsequent step,  $F'_{min}$  is calculated from  $F_1$  and  $F_2$  as indicated below (see Clark, 1973, p. 347):

$$F'_{min} = F_1 F_2 / (F_1 + F_2) \quad (6.1)$$

While  $F'_{min}$  represents a valid approximation to  $F'$  (Forster & Dickinson, 1976), some early replies to Clark's paper voiced concern about either of these two measures being overly conservative, arguing that Clark's approach unduly sacrifices power (the ability to detect effects where they exist) to protect against Type I errors (Wike & Church, 1976, e.g.). While the research community eventually followed Clark's argument for treating linguistic items as a random factor, with time the procedure to account for random subject and item variability markedly departed from Clark's solution to compute  $F'$  or  $F'_{min}$ . As discussed by Raaijmakers, Schrijnemakers, and Gremmen (1999), the use of  $F'/F'_{min}$  was gradually replaced by many researchers in favour of the ' $F_1 \times F_2$ ' criterion: following the reasoning behind this 'spin-off criterion', a treatment effect is considered significant if both  $F$  statistics of separate  $F_1$  and  $F_2$  ANOVAs turn out significant. While this  $F_1 \times F_2$  approach adopts the *first* step in the computation of  $F'_{min}$  in conducting separate by-subject and by-items analyses, it falls short of actually deriving a *joint* statistic from these.

As pointed out by Raaijmakers et al. (1999), however, the rationale behind the  $F_1 \times F_2$  criterion is flawed: it may well be the case that both  $F_1$  AND  $F_2$  are significant by themselves, while the respective  $F'_{min}$  computed from them is not. Raaijmakers et al. (1999) further note that part of the reason for this widespread adoption of the  $F_1 \times F_2$  criterion may stem from the above mentioned concern about the  $F'/F'_{min}$  procedure being too conservative as a test, a worry which Forster and Dickinson (1976) have shown is only warranted under some conditions. Thus, while Clark's paper provided a fix for the language-as-a-fixed-effect-fallacy

within the ANOVA framework, the recommended procedure became distorted in practice, with the resulting  $F_1 \times F_2$  approach failing to provide the required degree of generalisation.

### 6.1.3 Problems with ANOVA coping strategies: matching, factorisation, aggregation and correction

As discussed by Clark (1973) and Raaijmakers et al. (1999), it may not always be necessary for items to enter statistical analysis as a random effect, depending on the design of a given study. If items are nested within subjects – such that a different set of items is randomly sampled for *each* subject – or if the experimental stimuli are matched with respect to all potentially confounding variables across all treatment levels, conducting only a standard by-subject  $F_1$  ANOVA (and thus treating the language materials as a fixed effect) provides sufficient statistical evidence.

While matching of confounding variables is a common practice in experimental linguistics, an exhaustive matching approach involving an ever-growing number of potential candidate variables to be matched (ignoring currently *unknown* confounding variables) comes with its own problems: as Cutler (1981) and Amsel (2011a, chapter 4) point out, converging on an item set might turn out to be an almost insurmountable challenge or may result in the application of highly idiosyncratic and possibly unrepresentative materials, thus running counter to the general aim of ensuring generalisability of one's experimental findings to the relevant population of items.

Closely linked to the problem of matching covariates is the factorial design of experiments analysed with ANOVA (see Amsel, 2011a, chapter 4; Baayen, 2010; Cohen, 1983; MacCallum, Zhang, Preacher, & Rucker, 2002), i.e. treating all independent variables of the model as factorial variables with two or more *discrete* treatment levels. While some variables naturally fall into discrete classes (such as when testing the effect of word class membership), many other variables are naturally measured on a *continuous numeric* scale. However, making the latter fit into the ANOVA framework is usually achieved by dichotomising numeric variables into two (or more) discrete levels.

Baayen (2010) draws attention to the fact that dichotomisation may be an adequate measure for numeric variables which naturally lend themselves to being coerced to a treatment-coded variable, such as stimulus-onset asynchrony (SOA), where the effect may be known to be linear and which can be manipulated independently of intrinsic properties of linguistic items. For item-bound characteristics, such as word length, frequency or neighbourhood density, on the other hand, dichotomisation may entail a number of adverse consequences, including a considerable decrease of statistical power and effect size (Baayen, 2010; Cohen, 1983; MacCallum et al., 2002) or, conversely, under some circumstances overestimation of effect sizes along with spurious statistical significance when two or more dichotomised variables enter the analysis (MacCallum et al., 2002). In addition, by dichotomising one may run the risk of overlooking or misrepresenting non-linear effects (Baayen, 2010; MacCallum et al., 2002).

The computation of separate  $F_1/F_2$  ANOVAs requires aggregation of single trials over items and subjects, respectively, to retain by-subject and by-item condition means. As a consequence, a significant portion of the original variance present at the single trial level does not enter either analysis to begin with. Thus, the averaging procedure required entails a *loss of information* for statistical inference, which may again lead to overestimation of the size of effects (Locker, Hoffman, & Bovaird, 2007).

Other consequences of this information loss caused by aggregation over subjects/items also become apparent when trying to assess effects which *require* access to the original trial-level data set. Evaluating longitudinal effects which may arise in the course of an experiment (such as fatigue or learning effects) involves entering the trial number as a covariate into the statistical model. As exemplified by Baayen et al. (2008), inter-trial dependencies – such as spill-over from preceding trials in reaction time experiments – may be an important covariate warranting inclusion into one's model. Indeed, reaction times to preceding trials are often one of the best predictors for the response time on the actual target trial and controlling for this covariate within the statistical model may sometimes even resolve counterintuitive effects implied by analyses ignoring such dependencies (Baayen et al., 2008).

Baayen et al. (2008) discuss by-subject regression (or 'random regression') as one way of quantifying possible effects of continuous covariates while retaining the original trial-level structure: this approach first involves fitting separate regression models for each single participant, followed by assessment of significance of a given predictor by running one-sample *t*-tests on the set of coefficients for this predictor returned from the individual regression models. However, if these predictors gauge effects which are intrinsic to experimental items (e.g. word length or frequency), such an analysis will account for random subject variability, but it will fail to do so for random item variability where required (cf. Baayen et al., 2008).

Finally, another well known issue with ANOVAs lies in some of the a-priori assumptions about the data structure which have to be met for a valid analysis. For the analysis of repeated measures designs, application of univariate repeated measures ANOVA presumes *sphericity*, which holds when all the difference scores between treatment pairs have the same variance. In real life situations, this assumption is often violated, resulting in a too liberal test. When sphericity does not hold, one fix is provided by correction procedures such as the Huynh-Feldt and Greenhouse-Geisser corrections, which make the test more conservative by reducing the degrees of freedom for estimating the critical *F*-values. While such corrections provide an 'a-posteriori fix' for this issue, Quené and van den Bergh (2004) rightly stress that making use of an analysis which allows to estimate variances and covariances from the data is a more sensible approach than imposing a-priori assumptions (such as sphericity) on the data and subsequently correcting for violations of these postulated preconditions.

In brief, in this section I have reviewed a number of well known issues with ANOVA-based analysis of scalar data. While Clark's 1973 paper provided a way to deal with the language-as-a-fixed-effect-fallacy within the ANOVA framework, his original procedure has often been reinterpreted and watered down to application of the  $F_1 \times F_2$  criterion, undermining the general goal of ensuring generalisability of experimental findings to subject and item populations. Several authors have drawn attention to the pitfalls posed by exhaustive matching of

potentially confounding variables across treatment levels and dichotomisation of (multiple) continuous numeric predictors to fit the factorial nature of ANOVA-based designs, which may lead to unrepresentative stimulus sets, over- or underestimation of effect sizes and statistical significance. Further, aggregation prior to the separate  $F_1/F_2$  analyses entails a loss of information, by reducing the original amount of variance present at the single-trial level and by rendering assessment of inter-trial dependencies (e.g. effects of learning, fatigue or spill-over) impossible. Finally, ANOVAs are based on a-priori assumptions about the data structure – such as sphericity for repeated measures ANOVAs – which are frequently violated, forcing application of additional correction algorithms.

Despite these problems, the ANOVA approach to data analysis is still the predominant one in linguistics and related disciplines, such as cognitive psychology. The next section will discuss an alternative a tool which does not suffer from the particular issues listed above.

## 6.2 Linear mixed models as an alternative analysis tool

Linear mixed effects models (also known as '(linear) mixed models', 'hierarchical linear models' or 'multi-level models' – henceforth LMM) have gained considerable ground as an alternative to ANOVAs for data analysis in empirical linguistics (for reviews of LMMs in this research field see Baayen et al., 2008; Locker et al., 2007; Quené & van den Bergh, 2004). In part, this recent spread of LMMs is due to the increased availability and usability within commercial and open-source statistics platforms, coupled with a growing awareness of the fact that LMMs offer ways of tackling the ANOVA-related issues laid out in the previous section in a principled manner.

One of the key assets of LMMs is their ability to *handle multiple random effects within one and the same model* in a very flexible way: thus, both subject- and item-related random effects can be assessed simultaneously in a single step, without the need to conduct separate by-subjects and by-items analyses. In fact, any number of grouping factors whose levels require treatment as a random effect may be included. In addition, LMMs can handle a wide variety of experimental designs in a coherent manner: by defining more than one variance component, a LMM may account for multiple levels of variability and these levels may be *nested* within each other or *crossed* (Baayen et al., 2008; Bates, 2010).

Nested random effects arise in configurations of random factors where each level of one random factor is *exclusively* associated with *one* level of another random factor. One frequently cited example is the case of students, who are nested within classes, which are in turn nested within schools. As pointed out by Bates (2010), such strictly *hierarchical* clustering structures arise frequently in practice in a number of fields and explain the origin of the terms 'hierarchical' and 'multi-level'. However, LMMs are not limited to such settings, but can also handle crossed random effects, where *all* levels of one random factor occur with *all* levels of another random factor. Such a situation of course often holds in linguistic experiments with designs where each subject responds to the same item set, i.e. subjects and items are crossed random effects (Baayen et al., 2008; Barr et al., 2013; Bates, 2010).

In fact, LMMs can be applied in a rather straightforward way with numerous experimental designs commonly found in experimental or observational studies (see Baayen et al., 2008, for some designs frequently found psycholinguistic studies). Thus, LMMs offer a principled solution to the language-as-a-fixed-effect-fallacy by making it possible to include multiple random factors within one and the same model, enhancing generalisability of experimental findings beyond the immediate subject and item samples.

### 6.2.1 Flexibility

When using LMMs, the necessity to include a given variance component into the final model can be empirically assessed using *likelihood ratio tests*: by comparing models with and without the respective variance component against each other, researchers can test whether it is associated with sufficient variability to actually warrant inclusion. This procedure can also be applied to random correlation parameters, checking whether the final model should also account for covariance between random effects or not. Hence, using LMMs, investigators can flexibly specify the variance/covariance matrix of the model, which can be modelled on basis of the observed data, rather than making a-priori assumptions about the data such as imposing a sphericity condition for repeated measures designs (Quené & van den Bergh, 2004). In addition, rather than starting with a-priori assumptions about the (co)variance structure, Quené and van den Bergh point out that more can be gained by inspecting the estimated variance/covariance parameters, which often hold additional information about the data.

The term 'mixed effects' relates to the fact that LMMs can include a random effects part as well as a fixed effects part. While the random effects capture all relevant sources of random variability, the fixed effects specified indicate the estimated population average effects of the experimental variables of interest. Importantly, the fixed effects part can consist of categorical treatment variables, continuous numeric predictors as well as interactions between any of these two types of variables. The possibility to jointly estimate effects of categorical variables and continuous covariates provides a way for avoiding the pitfalls created by dichotomisation of numeric variables: rather than coercing continuous variables to variables with discrete levels, they can be included into the model as such, resulting in estimates of the associated slopes. This flexibility provides a way to avoid the risks associated with dichotomisation, such as under- or overestimation of effect sizes and statistical significance. Further, non-linear effects of continuous predictors may be captured using polynomials or splines, for example, affording insight into the presence of non-linearities (Baayen, 2010).

LMM analyses are based on the unaggregated data at the single-trial level, since averaging by subjects, items or other relevant grouping factors is not required. Consequently, the full amount of the original variance enters into the model, avoiding the loss of information entailed by averaging procedures (Locker et al., 2007). In combination with the option of making use of continuous predictors, access to the full trial-level data moreover naturally allows modelling of inter-trial dependencies such as longitudinal or spill-over effects (Baayen et al., 2008).

One more major advantage of LMMs over ANOVA frequently pointed out (e.g. Baayen et al.,

2008; Locker et al., 2007; Quené & van den Bergh, 2004) is the ability of mixed models to deal with unbalanced data sets exhibiting missing cases, provided that cases are missing at random. Since unbalanced data probably constitute the norm rather than the exception in real-world data analysis, this characteristic of LMMs is one of their major benefits.

Finally, LMMs may not only be applied for the analysis of continuous response data, such as reaction times, but they may also be used for the analysis of *binary* response data (e.g. correct/incorrect responses) by means of *generalized linear mixed models* with a logit link function. Thus, LMMs provide a general and consistent framework for modelling different types of data (Jaeger, 2008).

### 6.2.2 Open issues

The development of methods within the LMM framework currently is a very active area of research and while well-established procedures for parameter estimation exist, there remain a number of hotly debated issues. One of these relates to the computation of  $p$ -values for fixed effects of an LMM, the challenge being the determination of the degrees of freedom for their calculation, with different statistical software packages making use of distinct procedures (see Baayen et al., 2008, for discussion).

Also, there are currently no universally accepted standards for the best practice of defining mixed models. While LMMs afford the researcher to assess whether or not to include a given variance component in a 'data-driven' way, recent evidence based on simulation studies (Barr et al., 2013; Schielzeth & Forstmeier, 2009) suggests that it is crucial to include the *full* random effects structure implied by the experimental design to maximise generalisability. Barr et al. (2013) note that in many recent psycholinguistic studies relying on LMMs for data analysis, the random effects structures of the models ended up underspecified; while correctly accounting for random by-subject and by-item *intercepts*, many analyses fell short of including random *slopes* for the relevant random factors. By-subject intercepts adjust the overall intercept estimate for each individual subject, to the effect that noise will be reduced and sensitivity increased, since the random intercept variances are not confounded with any experimental effects of interest. Random by-subject slopes, on the other hand, account for between-subjects variability in their susceptibility to experimental manipulations.<sup>1</sup> Since the variances linked to random slopes are thus potentially confounded with the respective fixed effect of interest, leaving out a random slope which is actually implied by the experimental design increases the risk of Type I errors, thus undermining claims to generalisation of experimental results (Barr et al., 2013).

Barr et al. present simulation-based evidence suggesting that 'maximal' LMMs – i.e. models with the maximal random effects structure justified by the respective experimental design – show excellent performance in terms of staying close to the nominal Type I error level under different scenarios. Failure to include random slopes where indicated, however, results in

---

<sup>1</sup>The term random 'slope' may be somewhat confusing in being taken to refer only to continuous covariates. In fact, any kind of variable can enter the model as a random slope, whether factorial contrasts or continuous predictors.



inflated Type I error rates. Therefore, in settings involving confirmatory hypothesis testing, Barr et al. recommend always fitting models with the maximal random effects structure.<sup>2</sup>

While maximising the random effects part of an LMM may thus turn out to be the theoretically most appropriate approach to fitting mixed models, it can cause practical problems; for complex models with multiple random effects, the estimation algorithm may sometimes not converge. In such cases, it may be necessary to simplify the random effects part of the model, raising the question of which procedure to follow for such a simplification (see Barr et al., 2013, for suggestions). The problem of non-convergence is potentially aggravated with mixed logit models for binary outcomes, which may often render it infeasible to fit maximal models for non-continuous data.

### 6.2.3 Design structure

Despite these currently debated issues, LMMs have proven a viable and valuable alternative to ANOVAs, due to the numerous advantages addressed above. LMMs provide an efficient solution to the language-as-a-fixed-effect fallacy by allowing the inclusion of multiple random factors within one single model. Further, they allow to directly model the effects of continuous predictors, making dichotomisation unnecessary. By operating on the unaggregated single-trial level, LMMs make use of the full amount of information present in the data, whether in terms of the amount of variance entering the model or for the assessment of possible inter-trial effects.

For researchers, LMMs thus act as an analysis tool which provides a high degree of flexibility in accounting for the *design structure* underlying a given data set. By design structure, I will henceforth refer to all aspects of a data set which are relevant for its analysis and induced by the experimental design, including, for example, clustering caused by repeated measurements from subjects and items (or any other relevant random grouping factor) and inter-trial effects. While capturing a data set's design structure exhaustively is a crucial step in the analysis of any kind of data, doing so will still fall short of accounting for all important structural facets when we move from scalars to more complex kinds of data, such as curves.

## 6.3 Challenges of analysing curve data: towards holistic approaches

### 6.3.1 Holistic analyses: capturing design and object structure

Nowadays, curves are a frequently encountered data type in experimental linguistics and psychology and come in various guises. To begin with, I will focus on what we intuitively

---

<sup>2</sup>Note that random slopes are not limited to occur with the 'subject' random factor: Barr et al. (2013) urge to include by-item random slopes for *any* experimental variable which is *within* items. In fact, this principle should hold for any source of random variability – if an experimental manipulation of interest is *within-unit*, random slopes should be defined in addition to random intercepts.

conceive of as 'prototypical' curves, i.e. one-dimensional curves. Examples of well known relevant 1d-curve data are single-electrode ERPs, electromyographic curves representing, for example, activity of facial muscles, pupil dilation measured with eye-trackers or phonetic data such as electromagnetic articulography data and prosodic contours.

All of these examples refer to *temporal* signals, i.e. they represent signals dynamically evolving over time. While temporal signals are probably the most common one-dimensional curve data type observed in experimental linguistics, 1d-curves may also represent measurements along other domains. Taking an example from a very different area and methodology, curves acquired with mass spectrometry methods indicate the abundance of a molecule at a given mass-to-unit charge ratios ( $m/z$ ) value (see Crainiceanu, Caffo, & Morris, 2013; Morris, 2012; Morris et al., 2008, for reviews of analysis approaches for such data within the context of proteomic studies, for example). For the sake of simplicity, however, I will use temporal signals as examples for the remainder of this section. Note that while the number of data points in many kinds of linguistic 1d-curve data amounts to dozens or hundreds per sampled curve, other kinds of densely sampled 1d-curves can be high-dimensional, ranging into the tens or hundreds of thousands of data points for each curve.

Curves represent signals along some relevant biophysical axis (e.g. time) and, within the respective axis, adjacent or nearby measurements are usually correlated to some degree. Thus, such experimental curve data do not only exhibit design structure, but also an additional layer of relevant structure determined by the biological objects and processes generating the signals. In the remainder, I will refer to this extra level of structure as *object structure*, thereby referring to inherent internal structure of curves (or 'functions' – see section 6.5 below) of different kinds which should be acknowledged by the statistical analysis applied.

Ideally, any analysis of curve data should account for all sources of relevant structure, implying that a comprehensive analysis of curves should factor in the design structure *as well as* the object structure *jointly* – approaches to the analysis of curve data which meet this criterion will be termed 'holistic'. Note that the necessity of simultaneously accounting for these two levels of structure present with curve data is stressed by Ramsay and Silverman (2005, chapter 22), pointing out the importance of considering the aspects of *replication* and *regularization*, which in the present terminology correspond to design structure and object structure, respectively (see also Morris, 2015, for a recent discussion).

### 6.3.2 Standard approaches: failing holism (one way or the other)

ANOVAs are currently not only the standard for the analysis of scalar data, but also usually the method of choice for curve data. When used for analysing the latter, ANOVAs not only suffer from shortcomings on the design structural level, as discussed in the previous section, but also fail to account for inherent object structure. In the following, I will briefly review two strategies for ANOVA-based curve data analysis. Following Morris, Coombes, Koomen, Baggerly, and Kobayashi (2005) and Morris (2012), I will refer to the first one as 'feature extraction', which represents the most common procedure and involves a two-step procedure: first, temporal windows of interest are identified and then summary statistics

are extracted from the respective time windows. These summaries may involve averaging the values within a time window, or quantifying other features of interest, such as local 'peak' values (i.e. maxima or minima within a window) or latency information about local peaks. In a second step, these features are analysed using standard (uni- or multivariate) ANOVAs.

When sufficient a-priori information about the locus and nature of the effects of interest is available, focussing on a small number of time windows and features for analysis may suffice to identify relevant patterns in the data. Further, by considerably reducing the dimensionality of the potentially high-dimensional curve data, the feature extraction strategy certainly constitutes a computationally efficient and low-cost approach. However, feature extraction always involves coercion of curves to point data and a good part of the rich information contained in the data sets is often lost in the process, especially information about the exact extent and dynamics of possible effects. Depending on the underlying research goals, the loss may be negligible when the curves are of relatively short length or slow-wave curves, for example. But it may be substantial when high-dimensional curves with complex, fast changing and dynamically evolving patterns and many local features like peaks and spikes are the target of analysis. In such cases, the extent of the information loss entailed by feature extraction may be significant.

When assessing global effects by specifying one or few windows of interest which cover a large part of the curve, time-dependent properties of the relevant effects such as their exact onsets, offsets and maxima are lost. When defining narrow windows, approximate temporal localisation of an effect may be feasible, but details about its underlying dynamics are lost again. In addition, one may run the risk of misspecifying the extent of the windows, thus potentially missing an actually existing effect of interest by looking in the wrong place.

Alternatively, if inference about the detailed dynamics of effects as a function of time is of interest, one may apply the second strategy, a naive 'point-by-point' approach to curve analysis, which amounts to narrowing the analysis windows until their extent approximates or equals the original temporal resolution of the sampled curves. Possible effects are then assessed by conducting *independent* ANOVAs for each time point. Since such a procedure does not involve feature extraction on the object level, it reduces the risk of missing relevant effects and provides a way for tracing the dynamics underlying effects in detail over time.

Yet, such a point-by-point analysis only superficially exploits all of the object structural information: since the models are fitted independently for each time point, this procedure fails to capture correlations between adjacent and nearby data points and thus also fails to borrow strength for estimation across regions of the curves. Application of repeated measures ANOVA or MANOVA is no solution to this issue, since they still fail to take into account the autocorrelation between nearby points (see Levitin et al., 2007). By neglecting the correlational structure at the object level, point-by-point analyses also ignore the assumption that the processes generating the curves are more or less *smooth*.

On top of these shortcomings, due to the large number of pointwise tests, one will have to deal with a massive multiple comparisons issue and employing standard correction procedures (e.g. the Bonferroni correction to control for the family-wise error rate) may result in overly conservative tests, especially for high-dimensional data. Guthrie and Buchwald

(1991) present an alternative correction procedure for point-by-point testing in the context of ERPs, which takes the temporal autocorrelation of the data into account (see also Siegle, Ichikawa, & Steinhauer, 2008, for an application of this algorithm within a point-by-point approach to the analysis of pupil dilation data). However, rather than bringing in an essential characteristic of curve data in a post-modelling correction procedure, a more appropriate way of accounting for the correlational structure on the object level would be to make it part of the actual modelling process.

While the choice of the analysis model is independent of the question which of these two strategies to adopt, the feature extraction as well as point-by-point approaches to curve data often also involve prior averaging within subjects and/or items. In that case, the input to the final analysis consists of mean curves, rather than single-trial level data, and the analysis is subject to the caveats discussed in the previous section. While Bagiella, Sloan, and Heitjan (2000) advocate the use of LMMs for the analysis of psychophysiological (curve) data, they only cover the design domain, but do not address the question of how to best capture object structure.

Both of the above presented analysis approaches to curve data fall short of acknowledging the object domain structure. Thus, a more promising approach to a truly holistic analysis of curve data is one which tries to account for the intrinsic nature of curves right from the beginning. Before presenting one such framework, we will have a brief look of how these issues are tackled in standard and recent alternative approaches to the analysis of complex spatiotemporal electroencephalographic data.

## **6.4 Challenges of ERP data analysis**

### **6.4.1 ERP data: standard analyses approaches and their shortcomings**

ERPs are derived from the continuous electroencephalogram (EEG), which is an instance of a curve signal defined on a multi-dimensional space-time(-frequency) domain<sup>3</sup>. With modern equipment, the scalp EEG is typically sampled at high temporal rates (e.g. 500 Hz or higher) and may be recorded with high-density electrode setups involving between 64 and 256 electrodes, thus yielding quasi-continuous data in all dimensions.

While these multi-dimensional data sets hold rich information on each of the space and time (and frequency) axes, the canonical approach to data analysis involves discarding a significant portion of the original data: a typical ERP analysis involves averaging the single-trial segments cut around the relevant stimuli (or events) within subjects (and/or items) and conditions, resulting in mean ERPs for each of the experimental conditions. Analysis then proceeds by defining one or more temporal windows of interest (from tens to hundreds of milliseconds long) as well as spatial regions of interest (ROIs) encompassing one or multiple electrodes each. Once these parameters are set, feature extraction is performed by calculating either

---

<sup>3</sup> In the following, I will largely neglect the frequency domain – see chapter 9 for some discussion of possibilities for including the frequency axis into the modelling process.

mean values or features of local peaks (minimum/maximum amplitude values or their temporal latencies) for each time window and ROI. Typically, the extracted summary features are then entered as dependent variables into an ANOVA of appropriate type (cf. Dien & Santuzzi, 2005, for example).

Such standard ERP analyses thus proceed largely in parallel to standard analysis of 1d-curve data as discussed in the previous section, but necessarily also involve data reduction along the spatial object axis. This standard procedure does a great job at reducing the amount of the original data and thus minimising computational complexity required for the eventual statistical analysis, which has certainly contributed to its widespread use over the last decades. With sufficient a-priori knowledge and strong hypotheses about the temporal and spatial loci of effects of interest and careful selection of parameters, this approach captures relevant effects. The details about the underlying extent and dynamics, however, again get lost in the process and – now moving within a 3d-spatiotemporal object domain – the potential for information loss as well as the risk of missing potentially important patterns go up accordingly. In addition, the use of mean curves and ANOVAs also falls short of being able to account for the ‘fine structure’ on the design level.

While being an extreme case, the study of Thierry, Martin, Downing, and Pegna (2007) illustrates what may happen with improper specification of temporal windows/spatial regions of interest. Thierry et al. conducted an ERP experiment investigating the influence of physical interstimulus variance on face-selective aspects of the early N170 component. They concluded that larger N170 responses to faces than to pictures of objects are an artefact of physical low-level properties between the stimulus categories, a corollary with serious implications for a whole field of cognitive psychology. As discussed in detail in the critique by Rossion and Jacques (2008), however, the study of Thierry et al. suffered from a number of shortcomings and the lack of face-specific effects on the N170 can be mainly traced back to an inappropriate selection of electrodes for the ROI-based statistical analysis.

In addition to shrinking the complex object structure to a handful of scalar point data, the focus on features of local peaks often found in standard ERP analyses is conceptually unjustified, as stressed by a number of authors (e.g. Luck, 2005a, 2005b; Rousselet & Pernet, 2011), since peak latencies and amplitudes are confounded and “[t]here is nothing special about the point at which the voltage reaches a local maximum or minimum” (Luck, 2005b, p. 19).

Generally, canonical analyses of ERP-data involve extracting features from rich multi-dimensional data sets and feeding these to ANOVAs – they fall far short of being holistic approaches and further often concentrate on features of the data which do not have a special status from a conceptual point of view. The next section presents a number of recent alternatives which try to improve upon one or more aspects of standard approaches.

#### **6.4.2 Alternative approaches**

Rousselet and Pernet (2011) and Pernet, Sajda, and Rousselet (2011) plead for a change in the way ERP experiments are routinely planned and designed and their data are analysed.

Among other points, their agenda calls for single-trial level correlational analyses and for abandoning the focus on peak and time-window average measures (Rousselet & Pernet, 2011, p. 3): "Overall, there is no justification for limiting analyses to peaks or time windows of interest and throwing away the rest of the data. A systematic approach is thus necessary: analyzing all time points to reveal the complete time course of the effects." While this does not go the full way to a comprehensively holistic approach, these goals together make up an important subset of prerequisites and criteria for holism.

In psycholinguistics, some recent EEG- and MEG-studies made use of alternative statistical tools, covering different aspects of holism to different degrees: in an EEG study of word frequency and predictability on reading, Dambacher, Kliegl, Hofmann, and Jacobs (2006) applied repeated measures multiple regression analysis on the single-trial level to time window- and ROI-based means. Laszlo and Federmeier (2011) tested the effect of lexical and orthographical variables on N400 mean amplitude in visual word recognition with large-scale data set from 120 experimental participants. A multiple regression analysis was conducted using item-means of one single electrode and time window computed across participants; subject analyses were not conducted.

In an MEG study of lexical access in visual word processing, Solomyak and Marantz (2009) performed a ROI-based analysis within MEG source space, using single-trial regression at each time point for each ROI, normalising each subject's data to *z*-scores prior to analysis. In this study, corrections for multiple comparisons were conducted following the nonparametric cluster-based procedure introduced by Maris and Oostenveld (2007). Ettinger, Linzen, and Marantz (2014) used a variant of this approach using pointwise LMMs for the analysis of MEG data on auditory word processing.

Hauk et al. (2006) conducted an EEG experiment investigating the influence of a number of item-related properties on visual word processing, including word frequency and length as well as orthographical and conceptual covariates. In their analysis of the data, they went a step further than the above studies, by conducting separate analyses for each time point and electrode. The analysis applied was a two-level procedure, in which initial multiple regressions were done separately for each subject, followed by group-level statistical analysis using the estimates of the single subject regressions. The 'event-related regression coefficients' (ERRCs) obtained in this way as well as the corresponding *p*-values were presented as scalp maps of time-window mean values. Note that the *p*-values plotted in these maps were not corrected for multiple comparisons across time and/or space, based on the argument that the use of uncorrected *p*-values would make the result more directly comparable to previous studies making use the conventional statistical procedure. Hauk et al. (2009) closely followed this approach in another EEG experiment on visual word recognition.

Amsel (2011b) follows the same track analysing an EEG experiment on the effects of conceptual variables on visual word processing, using LMMs. In this study, standard LMMs were fitted at the single-trial level for each time point/electrode using by-subject and by-item random intercepts. Parallel to Hauk et al. (2006, 2009), results were plotted as scalp maps using uncorrected *p*-values for assessing statistical significance.

A number of recent studies have used the wavelet-based functional mixed model (WFMM)

of Morris and Carroll (2006), which will be discussed in detail below in section 6.6, for the analysis of EEG experiments. The WFMM permits fitting a 'functional' mixed-effects model to the complete curve data, without averaging within predefined time windows. Functional models are statistical tools tailor-made for the analysis of a wide variety of curve data which are able to capture the underlying object structure in flexible ways, going beyond naive point-by-point approaches; the functional data analysis framework will be presented more closely in the next section. The WFMM further allows to capture design structure flexibly by including multiple random effects within a model and allows to track a number of model parameters at each original time point, including fixed effects. Davidson (2009) made use of the WFMM in an analysis of data from a standard N400 experiment featuring semantic violations embedded within visually presented sentences using a single spatial ROI at the single-trial level. Subsequently, the WFMM was also applied in a study of morphosyntactic processing by Davidson et al. (2012), using two spatial ROIs.

In a similar spirit, Tremblay and Baayen (2010) and Kryuchkova, Tucker, Wurm, and Baayen (2012) applied generalised additive models (GAMs) (see Wood, 2006) in their analyses of ERP data. GAMs allow to jointly model the effect of multiple covariates over time, can capture non-linear effects of continuous covariates in a natural way and allow the inclusion of random effects. The single-trial model fitting was performed separately for each electrode and the complete segments were further split up into a series of overlapping smaller time windows to avoid losing precision in the smooths with longer time windows (Kryuchkova et al., 2012).

This brief survey of psycholinguistic studies deviating from the canonical ERP analysis involving averages and ANOVAs shows that the alternatives vary widely with respect to meeting holistic criteria. Conceptually closest to the general aim of minimising information loss are the WFMM- (Davidson, 2009; Davidson et al., 2012) and the GAM-based models of Tremblay and Baayen (2010) and Kryuchkova et al. (2012) as well as the 'naive' point-by-point approaches of Amsel (2011b) and Hauk et al. (2006, 2009). The latter are instances of the 'mass univariate' approach to multi- and high-dimensional data analysis popular in the analysis of functional magnetic resonance imaging (fMRI) data (Woolrich, Beckmann, Nichols, & Smith, 2009), which involves the fitting of separate models (of whatever type) to each voxel (for fMRI data) or time point/electrode combination (in the case of EEG-data).

A special challenge posed by such mass univariate analyses is the appropriate control of Type I errors, i.e. false discoveries, due to the large number of multiple comparisons, an issue explicitly ignored in the studies of Hauk et al. (2006, 2009) and Amsel (2011b) by the use of uncorrected  $p$ -values. In contrast, Hauk, Coutout, Holden, and Chen (2012) applied statistical parametric mapping (SPM; Friston et al., 1994) to the ERP data of another study on word recognition to model the detailed spatiotemporal extent of effects and used  $p$ -values corrected in a principled way by SPM's inbuilt method based on random field theory (Kiebel & Friston, 2004; Worsley, 2003). The assets and challenges of the mass univariate approach to EEG data analysis are discussed by Groppe, Urbach, and Kutas (2011), with a special focus on different recent appropriate techniques for correcting for multiple comparisons.

Note that the multiple comparisons issue does not only arise for point-by-point/mass univari-

ate analyses, but potentially also with approaches which try to account for the object-level structure in a more principled way, such as the WFMM. In section 7.7.2, I will present a novel approach to computing multiplicity-adjusted Bayesian posterior probabilities from the output of a WFMM.

Several software packages implementing some variant of the methods reviewed above are currently available: the SPM software used by Hauk et al. (2012) is widely used for fMRI analysis, but has been adapted for mass univariate spatiotemporal analysis of ERP data and is available as a MATLAB toolbox (Kiebel & Friston, 2004; Litvak et al., 2011). LIMO EEG (Pernet, Chauveau, Gaspar, & Rousselet, 2011) implements a closely related mass univariate approach, allowing hierarchical general linear modelling (GLM) of EEG data: in a first-level analysis, single-trial level data are modelled separately for each subject at each time point and electrode, followed by a second-level analysis which pools the parameter estimates obtained in the first-level analysis across subjects. As pointed out by Pernet, Chauveau, et al. (2011), this two-level approach is similar to that often used for the analysis of fMRI data as implemented, for example, in the SPM software. LIMO EEG further offers a number of different methods for multiple comparison corrections (e.g. using temporal or spatiotemporal clustering). It is implemented as a MATLAB toolbox and can be used via a graphical user interface. In a similar spirit, Smith and Kutas (2014a, 2014b) present the rERP (regression-based ERP) framework for regression-based mass univariate ERP analysis, which has been implemented as another MATLAB toolbox as well as in Python. It also offers a method for separating overlapping components arising from temporally close events using continuous EEG data, rather than epoched data (see also Burns, Bigdely-Shamlo, Smith, Kreutz-Delgado, & Makeig, 2013). The WFMM of Morris and Carroll (2006) is available as a standalone executable and is discussed in detail in section 6.6.5.

The brief outline of the various alternative approaches to ERP-data analysis presented here shows that they vary in meeting one or more criteria of holistic analyses and none of them covers them all. The remaining sections of this chapter will present the framework of functional data analysis as an alternative for curve data analysis and will introduce the WFMM as a specific model within this framework which goes a long way towards holistic modelling of curve data.

## **6.5 An alternative framework for the analysis of curve data: functional data analysis**

### **6.5.1 The general framework**

Functional data analysis (FDA) provides a set of tools tailored to the analysis of diverse types of curve data which are intended to capture the curves' inherent object structure. In their introduction to FDA, Levitin et al. (2007, pp. 135–136; emphasis in the original) provide a concise definition of the framework:



“A functional datum is not a single observation but rather a *set of measurements* along a continuum that, taken together, are to be regarded as a single entity, curve or image. Usually the continuum is time, and in this case the data are commonly called “longitudinal”. But any continuous domain is possible, and the domain may be multidimensional. In neuroimaging, for example, the activation levels observed at voxels in the brain are the responses over the two or three dimensions of space and possibly time as well. Or, data may be distributed over a continuous psycho-physical space (such as vowel space, musical pitch space, or colour space).”

The most conspicuous conceptual difference to the standard approaches sketched above is probably FDA’s focus on *curves as single entities* rather than treating them as a loose sequence of data points. The term ‘functional’ underscores this emphasis, implying that the curves are seen as *functions on a continuous domain*. In practice, each signal is of course measured at a finite number of sampling points, the ordered set of which I will refer to as *grid*  $t$  of length  $T$ . Thus, the sampled curves represent discrete approximations to the underlying continuous processes.

The above definition also shows that the scope of FDA goes beyond one-dimensional functions, covering functions defined on multi-dimensional domains, such as *2d-images* or *3d-volumes*. Note that I will be using the term ‘multi-dimensional’ to refer to the number of dimensions of the domain of a given functional data type (e.g. 1d-curves, 2d-images), whereas ‘high-dimensional’ refers to the number of grid points  $T$  of a functional data set. While multi-dimensional data are likely to also have high-dimensional grids, these two aspects of dimensionality are in principle independent and one-dimensional curve data may already be high-dimensional with tens or hundreds of thousands of sampling points per curve. Examples of 2d functional data encountered in linguistics and psychology include time-frequency images derived from single-electrode ERPs or ultrasound images of movements of articulatory organs (see Lancia, Rausch, & Morris, 2015, for the latter). Psychophysiological and neuroimaging methodologies yield numerous kinds of still more complex multi-dimensional curve data: structural brain scans can be considered 3d spatial *volumes*, while multi-electrode ERP and fMRI data represent multi-dimensional *spatiotemporal* functional data.

Levitin et al. (2007, pp. 138 ff.) provide a brief overview of the typical workflow of a functional data analysis, in which the first step usually involves applying nonparametric smoothing techniques to represent each sampled curve as a functional object. By making use of nonparametric smoothing approaches, a wide variety of curves can be smoothed without prior assumptions about their underlying shape. The smoothed curves thus estimated act as approximations to the underlying smooth process and are subsequently used for analysis, replacing the original raw data. Technically, the conversion to functional objects is commonly achieved by the use of *basis functions*, which serve as building blocks jointly representing smooth curves as *basis expansions*.

FDA methods make use of a repertoire of basis functions, which comprises pre-specified bases including Fourier series, splines or wavelets as well as empirically determined basis functions such as functional principal components (Crainiceanu et al., 2013; Levitin et al.,

2007; Ramsay & Silverman, 2005). The main factor in the choice of an appropriate basis is the nature of the signal to be analysed: Fourier series, for instance, can be used to model periodic signals, splines for complex, non-periodic curves and wavelets for complex, irregular and potentially (very) spiky signals.

Many of the statistical methods available for scalar data have been adapted to the functional setting, including functional ANOVA (*f*ANOVA) and functional linear mixed effects models (*f*LMM), for example. Functional linear models may feature functions as dependent variables (and scalar predictors), independent variables (and scalar outcomes) or functions on both sides of the equation (as in function-on-function regression), though for practical reasons I will focus on models involving functional responses and scalar predictors.

Levitin et al. (2007) and Crainiceanu et al. (2013) point out parallels between FDA techniques and classical multivariate methods, which both deal with data vectors. While a comparison is valid in many aspects, Crainiceanu et al. (2013) note the following crucial differences: in classical multivariate settings, the data vectors are markedly shorter than in the prototypical FDA case, since the latter can involve high-dimensional curves with hundreds of thousands of data points, whereas the former typically involve between two and a couple of dozen data entries. Further, as pointed out above, one key assumption behind FDA is that there exists a *smooth* process generating the discretely sampled curves – as Crainiceanu et al. (2013, p. 224) stress, “[t]his is fundamentally different from a vector that contains, for example, a subject’s body mass and index (BMI) and systolic blood pressure (SBP). Indeed, the functional process can always be sampled at a finer scale, whereas there are no ‘in-between’ points for BMI and SBP.” And crucially, with functional data the data points of each sampled curve are ordered in a natural way which reflects the intrinsic structure of the domain of the respective functional process. This ordering is crucial for analysing the data and interpreting the analysis’ results: thus, a measurement’s time stamp within a single-channel ERP curve or a voxel’s spatial position in three-dimensional structural brain scans are inherent characteristics of these data types.

Levitin et al. (2007) draw attention to the relevance of *derivatives* in FDA. The first derivative of a curve represents its velocity, while the second derivative indicates its acceleration, for example. Derivatives can not only be used to smooth the raw curves (e.g. by imposing roughness penalties at the level of the second derivative), but may provide additional information about the research question by unveiling the *dynamics* underlying the original (smoothed) curves. The exploitation of derivative information is currently an active area of research within FDA (cf. Ramsay & Silverman, 2005; Vines, Nuzzo, & Levitin, 2005).

FDA may be applied to achieve a number of different inferential goals, but it probably has one of its biggest advantages over more traditional analysis approaches when the aim is to trace effects as a function of time (or space, or some other domain) along the curve, by taking into account the function-internal covariance structure and by estimating the underlying smooth process. The gain potentially grows from a practical perspective with high-dimensional functions, since appropriately chosen basis functions may enable sparse representations of the functional objects, decreasing the effective dimensionality of the data while increasing parsimony and computational feasibility for big data settings and/or complex models.

While the FDA toolbox already contains a number of well-established procedures and models adapted to the functional setting, some aspects of functional modelling are still preliminary and the aim of ongoing research (Levitin et al., 2007; Liu & Guo, 2012), including suitable procedures for functional hypothesis testing and model selection as well as the question of how to account for multiplicity from the FDA point of view (for the latter, see Krivobokova et al., 2010, for example). Many FDA methods have been developed for the frequentist and Bayesian frameworks and different solutions for these issues are being developed within either. In the following chapter, I will present one novel Bayesian approach involving computation of multiplicity-adjusted posterior probabilities from simultaneous credible bands in more detail (see section 7.7.2).

Despite the abundance of functional data in linguistics, psychology and related fields, FDA methods are still a bit of an underdog, but are beginning to be applied to various kinds of data. Examples include Jackson and Sirois (2009) and Geangu, Hauf, Bhardwaj, and Bentz (2011), who collected pupil dilation data from infants to investigate early psychophysiological correlates of cognitive and emotional processing. In both studies, the functional data were represented by means of  $B$ -spline bases and time-dependent effects were assessed using subsequent ANOVAs. Aston, Chiou, and Evans (2010) applied a combination of functional principal component analysis followed by modelling of the associated principal component scores via standard LMMs for the analysis of linguistic pitch contours. Vines et al. (2005) made use of FDA to gain insight into listeners' time-varying emotional experiences during music pieces and demonstrate the potential of exploiting derivatives for assessing underlying dynamics. In addition, the WFMM of Morris and Carroll (2006) has recently been applied to different types of phonetic data such as measures of regularity of vocal fold vibrations (Lancia, Avelino, & Voigt, 2013) and 2d-ultrasound images of articulatory movements (Lancia et al., 2015).

Derivative information is the focus of Thivierge (2008), who investigated the effects of cross-modal integration on higher derivatives of ERP waveforms. Further ERP studies applying FDA methods include Bugli and Lambert (2006), who apply functional ANOVA with  $P$ -spline bases to study the effect of Lorazepam on the P300 using an auditory oddball experiment. Davidson (2009) first used the WFMM of Morris and Carroll for the analysis of ERP data from a standard N400 experiment featuring semantic violations embedded within visually presented sentences, followed by subsequent applications of the WFMM to ERP data presented by Steen (2010) and Davidson et al. (2012).

FDA is a growing field of relevance to numerous disciplines – Levitin et al. (2007) provide a compact overview of the aims and procedures of FDA from the viewpoint of cognitive psychology and Ramsay and Silverman (2005) present a wide variety of aspects of FDA, use cases and different functional models in depth.

### 6.5.2 Functional mixed effects models

While FDA techniques provide ways to account for the object structure of numerous types of curve data, models within the FDA framework are not automatically holistic approaches, since

not all of them also allow to capture the design structure with a maximal degree of flexibility. An *f*ANOVA, for example, is subject to the same limitations on the design-structural level as its scalar counterpart. *Functional* mixed models (FMMs), on the other hand, are natural extensions of LMMs for scalar data to the functional setting (Guo, 2002; Morris & Carroll, 2006), flexibly capturing the structure on both relevant levels – the design and the object domain – within one single model holistically (for recent reviews see Crainiceanu et al., 2013; Liu & Guo, 2012; Morris, 2015).

FMMs can exhibit a high degree of flexibility on either level: in the most general case, they allow the inclusion of fixed and random effects of any kind, thus inheriting the ability of LMMs to account for the design-induced dependency structure as well as inter-trial dependencies in a straightforward way. At the same time, FMMs can potentially capture the object structure of arbitrarily complex curves without prior assumptions on their form, whether for one- or multi-dimensional functional data. When the observations are functions and the predictors are scalars, such FMMs can provide functional estimates of the fixed and random effects, thus allowing to model both along the grid  $t$ . The choice of model is partly determined by the character of one's data: introducing an early version of a rather general FMM, Guo (2002) used smoothing splines for the functional fixed and random effects. Morris and Carroll (2006), on the other hand, equipped their variant of a FMM with a wavelet-basis, which can better represent very irregular curves with distinct local features such as spikes.

In presenting the general FMM framework in a more formal way, I will follow Morris (2012, p. 127). Given a sample of  $N$  curves  $Y_i(t), i = 1, \dots, N, t \in \mathcal{T}, \mathcal{T}$  a compact set, a general version of an FMM is given by equation 6.2:

$$Y_i(t) = \sum_{a=1}^p X_{ia} B_a(t) + \sum_{l=1}^m Z_{il} U_l(t) + E_i(t) \quad (6.2)$$

Here, a functional response  $Y_i(t)$  is linked to a set of scalar (categorical or continuous) predictors  $X_{ia}, a = 1, \dots, p$ , via *fixed effect functions*  $B_a(t)$ , which are of undefined form. Each of these functional fixed effects models the effect of its respective predictor across the whole domain of the function, e.g. time or space.

Correlation between functions can further be accounted for by including *random effect functions*  $U_l(t), l = 1, \dots, m$ , also of unspecified form, for which the random effects design matrix  $Z = \{Z_{il}\}$  reflects the data's design-induced clustering structure. Thus, these functional random effects can account for non-independence by repeated sampling from the same subject or other relevant experimental units. When required, several levels of variability can be specified for the random effect functions, introducing an index  $h = 1, \dots, H$ , such that the respective random effects design matrices become  $Z_h$ . These layers of variability may be in a nested or crossed relationship to each other, allowing to model a wide variety of experimental designs. Finally,  $E_i(t)$  indicates the residual error functions of unspecified form.

Another key component for completing the FMM specification are the assumptions about the covariance *between* and *within* the functions. A definition with very general scope is provided by Morris and Carroll (2006), who take the set of functional random effects  $U(t)$  to be a

realisation from a multivariate Gaussian process defined by the *between*-function covariance matrix  $P$  and the *within*-function  $T \times T$  covariance surface  $Q(t_1, t_2)$ , which can be denoted as  $U(t) \sim \mathcal{MG}\mathcal{P}(P, Q)$ . The error functions  $E(t)$  are assumed to be independent of the random effect functions  $U(t)$  and to follow  $E(t) \sim \mathcal{MG}\mathcal{P}(R, S)$ , with  $R$  the corresponding between-function covariance matrix and  $S$  the within-function covariance surface.

The same FMM can be expressed somewhat more compactly using discrete matrix notation as in equation 6.3, assuming that all curves are sampled on the same grid  $\mathbf{t}$  of length  $T$  (Morris & Carroll, 2006, p. 184):

$$Y = XB + ZU + E, \quad (6.3)$$

where  $Y$  represents an  $N \times T$  matrix of the  $N$  curves sampled on the grid  $\mathbf{t}$ .  $B$  is a  $p \times T$  matrix of fixed effects,  $U$  an  $m \times T$  matrix of random effects and  $E$  is a residual error matrix of size  $N \times T$ . The design matrix for the fixed effects curves is  $N \times p$  matrix  $X$  and  $Z$  represents the  $N \times m$  design matrix for the random effects curves. This model formulation essentially represents the functional analogue of the LMM applied for the analysis of scalar data. The random effects  $U$  follow a matrix normal distribution  $U \sim \mathcal{MN}(P, Q)$ ,  $P$  representing the  $m \times m$  between-row covariance matrix and  $Q$  the within-row  $T \times T$  covariance matrix.  $E$  is assumed to be  $\mathcal{MN}(R, S)$ , with  $R$  and  $S$  of dimensions  $N \times N$  and  $T \times T$ , respectively. The within-curve covariance matrices  $Q$  and  $S$  here become discrete approximations to the continuous covariance surfaces.

While it is most natural to conceive of model 6.3 being applied to one-dimensional curve data, it can be generalised to the multi-dimensional case in a rather straightforward way (Crainiceanu et al., 2013; Martinez, Bohn, Carroll, & Morris, 2013; Morris, 2012, 2015; Morris et al., 2011). 2d-image data, for example, can be modelled by transforming each of the observed images into a row vector of length  $T = T_1 \times T_2$ , such that the resulting rows can be stacked to yield the  $N \times T$  data matrix  $Y$ , whose columns index pixels within the images. Applying the same logic to  $B, U, E, Q$  and  $S$ , the model accounts for all quantities in 2d-space. The same procedure can be applied to model functional data on a domain with three or more dimensions. In all cases, the autocovariance should be modelled in a way which does not only take into account one-dimensional proximity in the vectorised data, but the proximity in all relevant dimensions (e.g. proximity in the horizontal, vertical and diagonal directions for 2d-images).

FMMs thus offer a comprehensive solution to the challenge of accounting for the design and object structure underlying different kinds of functional data in one unified, holistic model in a maximally flexible way (see Crainiceanu et al., 2013; Liu & Guo, 2012; Morris, 2015, for recent reviews). Though FMMs are a relatively recent innovation, different FMM-variants have already been used in the analysis of various kinds of functional data from a number of fields such as proteomics, neuroscience or linguistics (Crainiceanu et al., 2013).

As remarked by Liu and Guo (2012), ultimately the success of FMMs will depend on the availability of efficient and user-friendly software implementations. In the development of these, one faces a number of non-trivial challenges posed by the high degree of flexibility

afforded by FMMs and the potentially (extremely) high-dimensional nature of the data. With data sets comprising hundreds of millions or even billions of data points, even loading the entire data matrix into memory in one step will not be possible on standard computers. Another source of computational complexity are the covariance matrices and handling these can already become an issue with data of more modest dimensionality.

Thus, *big data* settings require the combination of a number of different strategies, including dimensionality reduction via functional principal component analysis, subsampling of the data and the development of sequential fitting procedures which do not require access to the full data set at a time (see Crainiceanu et al., 2013, section 13.4.5). Other key aspects here are the choice of basis function and the concrete approach to model fitting taken – a suitable choice of both can lead to sparse representations of functional objects and a combination of flexible yet parsimonious modelling (Crainiceanu et al., 2013; Morris et al., 2011).

Depending on the dimensionality of the data set, the complexity of the model to be fitted and the soft- and hardware used, computational feasibility may become an issue in practice and one may have to compromise. Also, not all of the FMMs which have been proposed have been implemented as readily available software and those which have may be subject to a number of inherent and practical limitations (Liu & Guo, 2012). While FMMs promise to act as holistic modelling tools providing a maximal degree of flexibility for analysing complex functional data, their actual impact will also depend on their ability to scale up to high- and multi-dimensional data sets and the availability of accessible and user-friendly software. While different solutions to these challenges are currently explored (Crainiceanu et al., 2013), the next section will take a closer look at one specific method, the WFMM of Morris and Carroll (2006).

## 6.6 The wavelet-based functional mixed model

### 6.6.1 Model definition

The wavelet-based functional mixed model developed by Morris and Carroll (2006) combines a functional mixed model approach with wavelet basis functions and embeds the modelling in a fully Bayesian framework. By combining these features, the WFMM represents a versatile tool for the analysis of a wide range of curve data.

The basic structure of the WFMM is given by the general model 6.2 (see equation 6.3 for the same model in matrix notation) and complemented by the presentation of the assumed covariance structure laid out in the previous section. Accounting for its ability to handle  $H$  levels of random variability, a compact representation is given by model 6.4:

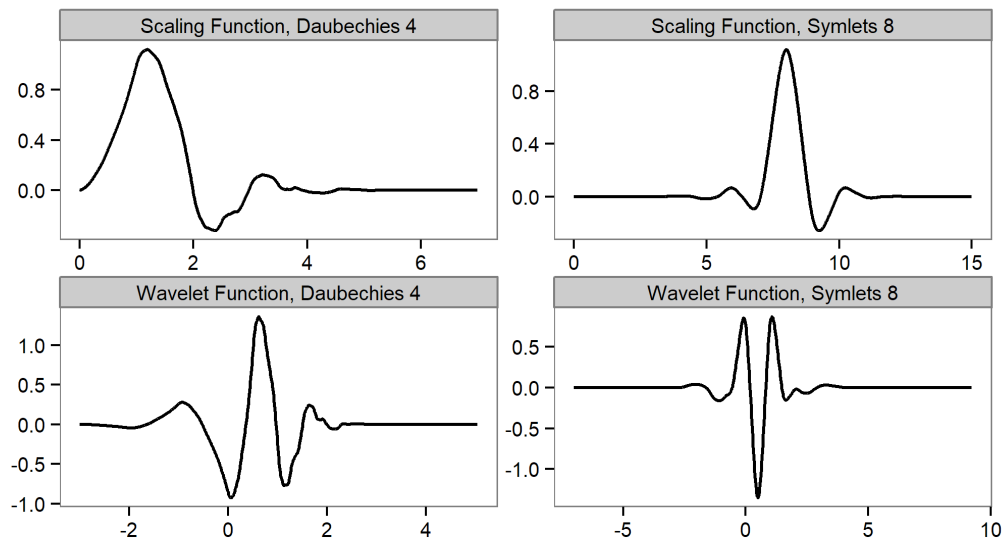
$$Y_i(t) = \sum_{a=1}^p X_{ia} B_a(t) + \sum_{h=1}^H \sum_{l=1}^{m_h} Z_{il}^h U_l^h(t) + E_i(t), \quad (6.4)$$

with  $U_h(t) \sim \mathcal{MG}\mathcal{P}(P_h, Q_h)$  and error functions  $E(t) \sim \mathcal{MG}\mathcal{P}(R, S)$  independent of the

random effect functions  $U(t)$ .

## 6.6.2 Wavelets: a quick tour

While this specification already provides a great deal of freedom for modelling a wide range of experimental designs, another crucial source of flexibility on the object structural level is found in the WFMM's use of *wavelet basis functions*. The following review of some of the key features of wavelets is largely based upon the accessible conceptual review offered by Samar, Bopardikar, Rao, and Swartz (1999).



**Figure 6.1:** Scaling (top row) and wavelet (bottom row) functions of the 'Daubechies 4' (left column) and 'Symlets 8' (right column) wavelets.

Wavelets are oscillating functions of time<sup>4</sup> which have a number of key properties making them widely used basis functions in signal analysis. A key feature of wavelets is their ability to represent non-stationary signals of arbitrary form. Wavelet functions come in a multitude of different shapes, each belonging to a wavelet 'family' of basic shapes. Figure 6.1 exemplifies the wavelets functions of a *Daubechies 4* ('Daubechies' family) and *Symlets 8* (a.k.a. 'least asymmetric' Daubechies wavelet – 'Symlets' family) wavelet and its associated scaling functions.<sup>5</sup>

One of the characteristics of wavelets which makes them attractive for a wide range of applications is their relative *localisation in both time AND frequency*: while sine and cosine waves

<sup>4</sup>For the sake of illustration, I will use one-dimensional time series, though wavelets can also be used for decompositions of signals on other (multi-dimensional) domains – see below.

<sup>5</sup>All plots within this section are based on output of functions of the *wavethresh* package for R (Nason, 2013).

are precisely localised in frequency, they extend infinitely on the time domain. Wavelets, on the other hand, can be *compactly supported* in time, i.e. a wavelet's 'energy' is restricted to a very specific interval of time. Wavelets are also *band-limited*, i.e. a wavelet transform decomposes a signal at a number of frequencies ('scales') – crucially, a given wavelet will only be able to 'see' and pick up that part of the signal which lies within its respective time and frequency range.

The two key operations performed on wavelets to decompose signals simultaneously in time and frequency are *scaling* and *translation*. Scaling involves 'shrinking' and 'stretching' wavelets along the time axis: *stretching* a wavelet will shift the part of the frequency spectrum covered to *lower frequency* scales and its spectrum will concentrate over a smaller bandwidth, thus making it *more localised in frequency* – at the same time, since a stretched wavelet will cover a longer time range, it will be *less localised in time*. *Shrinking* a wavelet, on the other hand, will make it cover *higher frequencies*, making it *less localised in frequency*, since the bandwidth covered will increase, but *more localised in time*. Translating a wavelet amounts to *shifting* it to different time positions at any scale.

By performing a wavelet analysis of a signal, it is thus "divided into a set of smaller waveform pieces that capture information selectively about the waveform's structure at specific times and within specific frequency ranges" (Samar et al., 1999, p. 15). The scaled wavelets are shifted along time at a given scale and with each translation the correlation of the wavelet's shape with the part of the original signal which is 'in view' of this particular wavelet is computed.

Wavelets on larger scales (i.e. lower frequency ranges) will pick up the large-scale fluctuations of the signal being decomposed, while smaller-scale wavelets will only correlate with the higher-frequency details contained in the curve. The correlations thus calculated are the *wavelet coefficients*. The smaller the scale of the respective wavelet, the higher the level of detail a corresponding wavelet coefficient will represent – this way, a wavelet transform decomposes a signal at various levels of resolution, allowing to assess its underlying large-scale patterns or to zoom in to look at the small-scale details with a single analysis step.

A wavelet analysis can either be conducted using the *continuous* or the *discrete wavelet transform* (CWT and DWT, respectively): the CWT involves scaling and translating a wavelet in very small steps, resulting in highly redundant representations. For the DWT, only a special subset of possible scalings and translations is performed and the resulting set of wavelet coefficients is a non-redundant representation of the original signal, where the number of coefficients equals the number of sampling points of the original signal.

Crucially, the coefficients returned by a DWT are *orthogonal*, i.e. *the coefficients representing the different scalings and translations used in a DWT do not correlate with each other* within and across scales and translations. Thus, the DWT returns an orthogonal basis representation of the original signal. Despite its non-redundancy, this representation is lossless and the original waveform can be reconstructed by applying the *inverse discrete wavelet transform* (IDWT).



The DWT is computed by a simple recursive filter scheme involving the iterative application of a high- and low-pass filter followed by downsampling the resulting wavelet coefficients by a factor of two. In each iteration, the high-pass filtering capturing the small scale (high-frequency) details is performed using the wavelet functions (see figure 6.1, bottom row), while the low-pass filtering resulting in wavelet coefficients which represent the large-scale (low-frequency) patterns is done by the associated scaling functions (figure 6.1, top row).<sup>6</sup> To reconstruct a signal from the resulting wavelet coefficients, the procedure is reversed. Importantly, for both the DWT and the IDWT fast and efficient computational algorithms exist.

Figure 6.2 shows the set of wavelet coefficients resulting from the DWT of a single curve containing 512 sampling points ( $T = 512$ ). The original curve is the result of a random process and is shown in the bottom panel; the wavelet used for decomposition was the *Symlets 8* illustrated in figure 6.1 with 8 levels of decomposition. The DWT coefficients are plotted on top of the original signal, increasing level numbers index higher-scale wavelets and the position on the x-axis represent the translation within each scale. Application of the IDWT to these wavelet coefficients will reconstruct the original signal. Note that the ranges of the y-axes vary across levels and their extent tends to decrease for higher levels – bearing this in mind, it is easy to see that most of the wavelet coefficients representing high-frequency details (e.g. levels 6-8) are actually close to zero.

In many settings, the most important features of a signal indeed tend to be represented by a small number of wavelet coefficients only, while white noise is distributed over all of the coefficients equally. This property of wavelets is particularly useful in applied settings, since it allows to represent signals with *sparse* bases, for example by setting all coefficients below a computed threshold to zero. By applying such a *wavelet compression*, the originally noisy curves can also be *denoised*. The degree of fidelity to the original signal depends on the algorithm used and the parameters chosen.

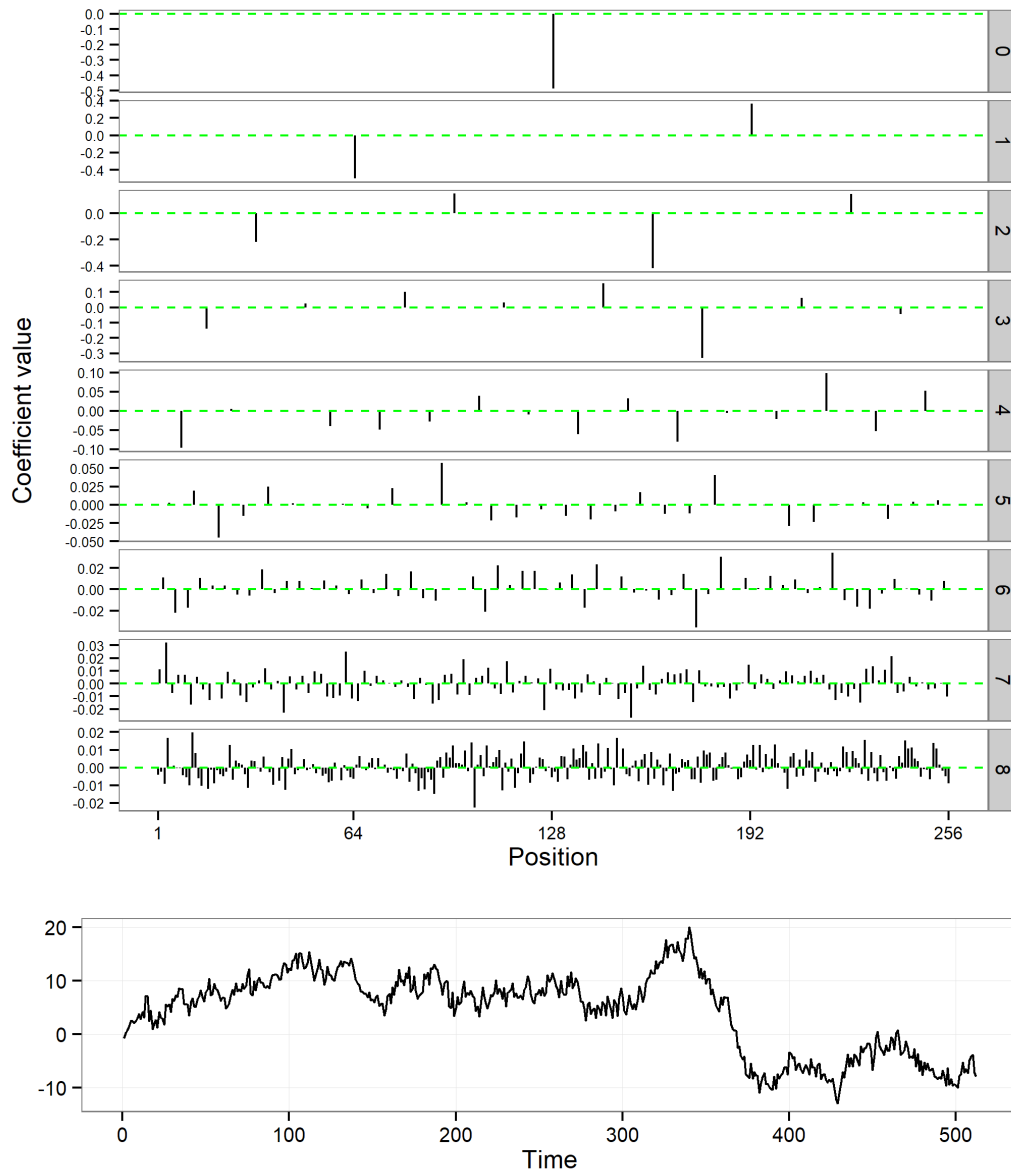
Figure 6.3 illustrates the result of thresholding the coefficients of figure 6.2 at levels 4 to 8. The result is a considerably thinned out representation, retaining only 31 coefficients, which amounts to 6.1% of the original number of coefficients. The curve reconstructed from this highly reduced set of coefficients is a smoothed version of the original. Note that all of the original 256 coefficients of the highest detail level (8) have been set to zero and of the highest 3 levels (6-8), only 3 out of originally 448 coefficients are retained in the compressed version.

Figure 6.4 shows the result of limiting the same thresholding operation to level 8 only, setting all coefficients at this level to zero and leaving 255 coefficients, (49.9% of the original amount). In the reconstructed curve, only the noise at the highest frequency range is cancelled out and the result is much closer to the original.

Thus, wavelets have a number of characteristics which makes them an excellent choice of basis functions for a wide range of signals: they are extremely flexible and can represent smooth trends as well as abrupt changes and local spikes at the same time and can be used

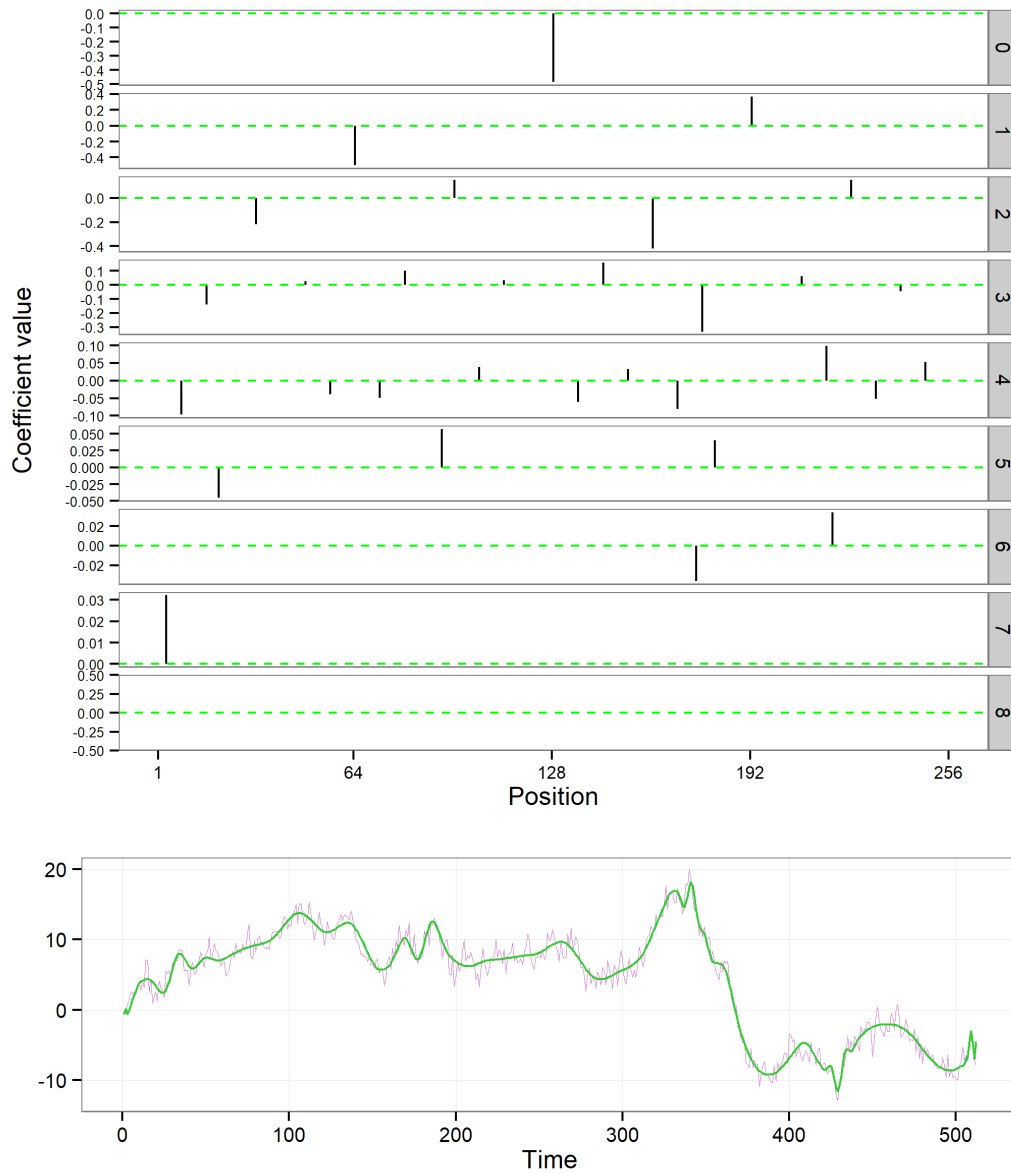
---

<sup>6</sup>For more details, see Samar et al. (1999).



**Figure 6.2:** Bottom panel: simulated curve. Top panel: wavelet coefficients for the DWT of the curve at each level of decomposition. Higher level numbers indicate higher levels of detail, i.e. higher frequencies.

to model non-stationary curves. Two further key features of the DWT which allow parsimonious modelling of complex and potentially high-dimensional curves are the orthogonality of the resulting set of wavelet coefficients and the ability to obtain sparse and denoised representations of functional objects through wavelet compression.

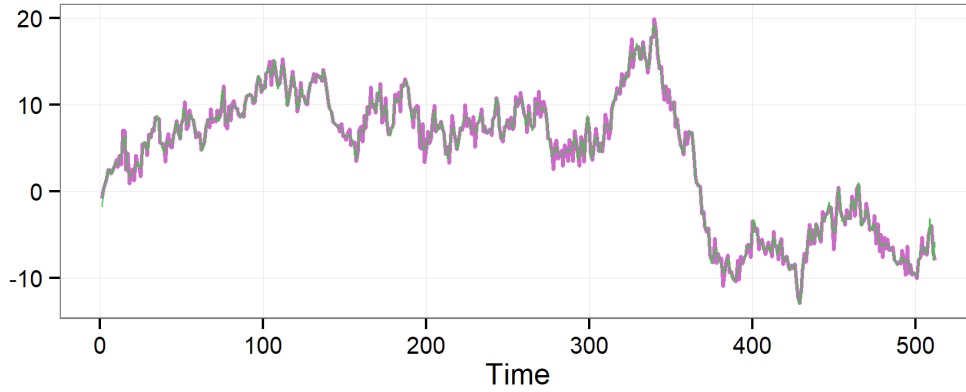


**Figure 6.3:** Bottom panel: original (purple) and compressed curve (green). Top panel: thresholded DWT coefficients for the compressed curve at each level of decomposition, with compression performed at levels 4–8.

### 6.6.3 Internal WFMM architecture, modelling processes and scope

Fitting a WFMM involves a basic three-step procedure (cf. Morris & Carroll, 2006):

1. The sampled functions are transformed to the wavelet space through a DWT.



**Figure 6.4:** Original (purple) and compressed curve (green). Here, only wavelet coefficients at level 8 were thresholded.

2. *Model fitting is performed in the wavelet space*, using the DWT coefficients. Estimation of all parameters is done within the Bayesian framework using Markov chain Monte Carlo (MCMC) sampling. The result is a set of MCMC samples for a number of functional model parameters in wavelet space.
3. The wavelet space MCMC samples of a number of parameters are projected back to the data space using the IDWT, where the results can be easily interpreted and where any kind of Bayesian inference can be conducted.

Thus, in the initial step in fitting a WFMM, each row of the  $N \times T$  data matrix  $Y$  is transformed to the wavelet space by applying the DWT into a vector of  $T$  wavelet coefficients, yielding the wavelet space  $N \times T$  data matrix  $D$ . Each row  $i$  of  $D$  then holds the wavelet coefficients representing curve  $i$  of  $Y$ . While the columns of  $Y$  are indexed by the position/timestamp  $t$  within the original curve, the columns of wavelet space matrix  $D$  are double indexed by scale  $j$  and location (translation)  $k$ . The wavelet space model of the WFMM (Morris & Carroll, 2006, pp. 184–185) can be written as

$$D = XB^* + ZU^* + E^*, \quad (6.5)$$

where  $X$  and  $Z$  are the design matrices for the fixed and random effects, as in model 6.3. The rows of  $B^*$ ,  $U^*$  and  $E^*$  represent the DWT of the rows of  $B$ ,  $U$  and  $E$ , with all columns again indexed by scale  $j$  and location  $k$ . The distributional assumptions are  $U \sim \mathcal{MN}(P, Q^*)$  and  $E \sim \mathcal{MN}(R, S^*)$ , with  $R^*$  and  $S^*$  corresponding to the wavelet space versions of the within-function covariance matrices  $R$  and  $S$ . Crucially, due to the orthogonality of the DWT coefficients,  $R^*$  and  $S^*$  become diagonal matrices,  $Q^* = \text{diag}(q_{jk}^*)$  and  $S^* = \text{diag}(s_{jk}^*)$ .

The diagonality assumption on  $R^*$  and  $S^*$  reduces the number covariance parameters for each from  $T(T + 1)/2$  to  $T$  – given that  $T = 500$ , for example, only 500 instead of 125250

parameters for the fully unstructured variants are required. Since estimation of the covariance matrices can quickly become a computational bottleneck (especially with high-dimensional data), this diagonalisation of the covariance matrices is one of the major benefits gained by performing the modelling in the transformed wavelet space.

As pointed out and illustrated by Morris and Carroll (2006), independence in the wavelet space does not necessarily imply independence in the data space, however: larger variances at scales representing low-frequencies will indicate stronger serial correlations and smoother curves. Since all wavelet space model parameters are indexed by both scale  $j$  and location  $k$ , the WFMM does not only handle non-stationary features such as varying degrees of smoothness in the fixed effects functions, but also in the random effects and error functions. Thus, the model can accommodate any combination of smooth or spiky fixed effects functions and spiky/smooth random effect functions.

Here, the advantages of combining the wavelet basis with what Morris et al. (2011) term *isomorphic modelling* come into play. Isomorphic modelling refers to a generalisation of the basic three-step procedure the WFMM follows; it involves a *lossless* transformation of the original data (or near-lossless transformation; see Meyer et al., 2015), taking advantage of specific properties of the resulting basis objects by modelling the obtained coefficients in the transformed space and back-projecting the estimates to the data space. Often, the modelling can be done in a more parsimonious manner in the transformed, rather the original data space, thus increasing computational efficiency and feasibility of the modelling process while retaining a high degree of flexibility. At the same time, the results can be assessed in the data space, where interpretation is straightforward and intuitive.

The Bayesian approach to model fitting pursued in the WFMM further allows to implement a number of additional key modelling steps in a natural and unified way. Since an overview of the Bayesian framework goes far beyond the scope of the current work, in the following I will presuppose some basic knowledge about key aspects of the Bayesian approach to data analysis (see Kruschke, 2010a, 2010b, 2010c, for accessible introductions to the Bayesian approach from the point of view of cognitive psychology). For readers not familiar with Bayesian data analysis, I will nevertheless try to work out the conceptual benefits of specific WFMM features for estimation and inference.

While vague proper priors are placed on the WFMM wavelet space variance components  $q_{jk}^*$  and  $s_{jk}^*$ , a *spike-and-slab* mixture prior is used on the wavelet space fixed effects  $B_{ijk}^*$ , where  $i$  represents the  $i^{th}$  fixed effect function,  $B_{ijk}^* = \gamma_{ijk}^* \mathcal{N}(0, \tau_{ijk}) + (1 - \gamma_{ijk}^*) I_0$  and  $\gamma_{ijk}^* = \text{Bernoulli}(\pi_{ij})$  (Morris & Carroll, 2006, section 4.3), with  $\pi_{ij}$  and  $\tau_{ijk}$  representing regularisation parameters which can be estimated from the data using an empirical Bayes procedure. Less formally, the spike-and-slab prior consists of a sharp spike placed on zero and a broad normal distribution around zero (the 'slab' part).

The practical upshot of using such a prior is that each fixed effect function is *adaptively regularised*: it smooths the fixed effect functions by inducing non-linear 'shrinkage' and thresholding on their wavelet coefficients in such a way that prominent local peaks are preserved, i.e. the degree of smoothness can vary quite flexibly along  $t$  as indicated by the data. Since each fixed effect function  $i$  is adaptively regularised separately, they can have different global and

local smoothness properties. The Bayesian approach to model fitting here offers a natural way of performing locally adaptive smoothing by using appropriate priors.

Adaptive regularisation of fixed effect curves may be complemented by wavelet compression, along the lines introduced in the previous section. Compression of the DWT coefficients may reduce the full set of coefficients significantly, reducing computational workload and speeding up the modelling process. The WFMM software offers an option to conduct joint wavelet compression in a way which preserves a predetermined amount of energy and excludes wavelet coefficients insignificant across all curves.

The WFMM has been generalised in a number of ways: Morris et al. (2011) extended the WFMM to the multi-dimensional setting for modelling 2d-images and 3d-volumes by using multi-dimensional wavelet transforms and describe the possibility of applying other kinds of basis functions within the general framework of isomorphic functional mixed models. Zhu et al. (2011) implemented a robust version of the WFMM which down-weights outlying curves or parts of curves, resulting in improved precision of the estimates, and Zhu, Brown, and Morris (2012) showed how to perform classification of functional data with the robust and Gaussian versions of the WFMM.

Early applications of the WFMM include mass spectrometry proteomics data (Morris, Brown, Baggerly, & Coombes, 2006; Morris et al., 2008) and accelerometer data representing school children's physical activity (Morris, Arroyo, Coull, Ryan, & Gortmaker, 2006). Subsequently, the WFMM has been applied to a number of different kinds of data from linguistics and related fields: Davidson (2009) presents a WFMM of ERP data from an N400 sentence processing experiment (see also Steen (2010) as well as Davidson et al. (2012) for further applications of the WFMM to ERP data). WFMM-based analyses of phonetic data include measures of regularity of vocal fold vibrations (Lancia et al., 2013) for 1d-curve data as well as spectrograms of bat chirp syllables (Martinez et al., 2013) and automatic quantitative analysis of ultrasound images of the movements of articulatory organs (Lancia et al., 2015) for 2d-image data.

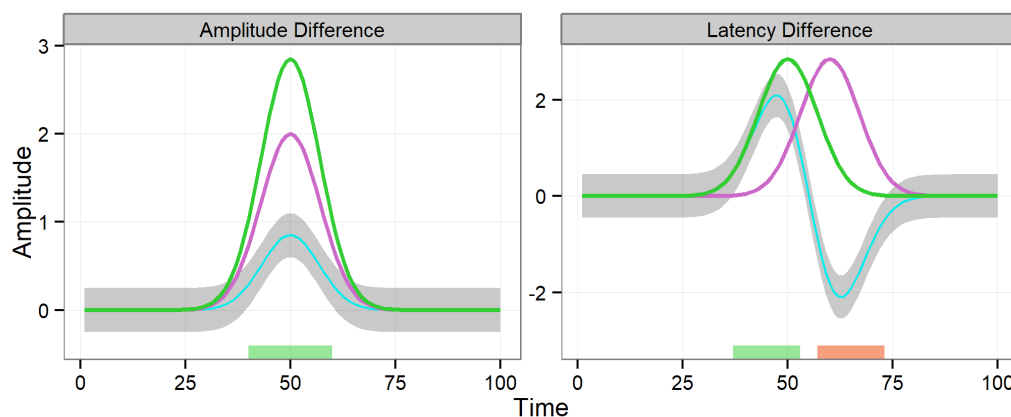
Thus, the diversity of the data falling within the scope of the WFMM underscores its versatility and its potential as a tool for holistic analysis of linguistic and psychological curve data. In practice, however, application of the WFMM to a given data set may be subject to a number of limitations. After a quick look at inference within the WFMM framework, I will take a closer look on the software implementing the WFMM as well as its strengths and current weaknesses in section 6.6.5 below.

#### **6.6.4 Inference in the WFMM**

One more benefit of the Bayesian approach to fitting a WFMM is the ease with which inference on the resulting model estimates can be performed; once the posterior samples for the parameters of a WFMM are available in the data space, it is straightforward to compute Bayesian credible intervals for each of the posterior fixed effect curves, for example. These may either be simple pointwise credible intervals based on quantiles or highest posterior

densities, simultaneous/joint credible bands which attempt to take into account the multiplicity along  $t$  (e.g. Krivobokova et al., 2010; Steen, 2010), posterior pointwise probabilities or variants which jointly account for multiplicity and practical significance (Morris et al., 2008) or any other Bayesian inference of interest (see section 7.7.2 below for novel approaches for computing different types of multiplicity-adjusting '*simultaneous probabilities*' from simultaneous credible bands, such as the '*simultaneous band scores*' (SIMBAS) of Meyer et al., 2015).

Figure 6.5 illustrates how to interpret graphical summaries of posterior fixed effect curves: the left panel contains two curves representing hypothetical data-based mean curves of two levels of an arbitrary binary variable. The turquoise line indicates the functional posterior mean estimate for the difference between the green (reference level) and the purple curve at each  $t$  (i.e. the most credible parameter value for each time point) and the grey band is the corresponding 95% credible interval for the parameter estimate. When basing inference on credible intervals of fixed effect curves, any  $t$  at which the interval does not contain zero is flagged as significant. The pattern in figure 6.5 suggests that there is a credible difference in *amplitude* between the two curves between time 40 and 60, indicated by the green bar at the bottom of the panel.



**Figure 6.5:** Left: the credible interval indicates an amplitude difference between the green and purple curves from time 40 to 60 (green bar at the bottom). Right: the pulse-like shape of the credible interval points to a difference in latency between the two curves (the green bar first indicates a positive-slope effect, followed by the negative-slope effect marked by the orange bar).

Note that a WFMM analysis not only provides information on amplitude differences, but allows to get inference about amplitude and latency effects (see Morris et al., 2008, 485–486). The right panel of figure 6.5 again shows two mean curves – this time, they are identical in amplitude, but the purple one is shifted along the time axis by ten units. Here, the corresponding functional fixed effect estimate follows a pulse-like pattern, first showing a positive difference and then a negative one. Thus, the shape of the effects may make a difference in interpretation and pulses may point to systematic effects of a predictor on the

latency of a local peak.

While a standard feature extraction approach would require two separate analyses to assess effects on amplitude and latency, the functional nature of the WFMM in principle allows to conduct holistic inference on these two parameters in one step. More information on basing inference about fixed effects on alternative indicators such as simultaneous credible bands or probabilities can be found in sections 7.6.2.4 and 7.7 in the following chapter.

### 6.6.5 The software: features and limitations

As rightly pointed out by Liu and Guo (2012), for FMMs to become widely accessible to researchers, they have to be available as user-friendly software. The WFMM of Morris and Carroll has been implemented as freely available standalone executable (see chapter 7). Generally, the software provides a great deal of flexibility, while at the same time minimising the amount of input required to be supplied by the user. Thus, the only obligatory input lies in the data matrix  $Y$ , the fixed effects design matrix  $X$  as well as the random effects design matrix  $Z$ . If nothing else is provided, it will run with default settings for all options, such as the settings for the wavelet basis, the MCMC procedure and the post-processing of the posterior samples. At the same time, users can easily modify these and more settings if desired. The program returns the raw posterior samples of a number of model parameters as well as summaries of the most important posterior quantities. The WFMM software also considerably reduces the user's workload by setting up default priors and estimating the regularisation parameters automatically from the data.

While the input required from the user is minimised, assembling the  $X$  and  $Z$  design matrices is not a trivial task for many researchers, since this is usually accomplished via high-level interface elements. When specifying a LMM with R's *lme4* package (Bates, Maechler, & Bolker, 2012), for example, the user defines the model via a model formula and the design matrices are automatically generated internally by lower-level functions. Also, when working with such a high-level interface, a model can easily be parametrised in different ways by recoding variables. Thus, many users may not feel comfortable when faced with the challenge of constructing these matrices by themselves.

Potential new WFMM users may further be faced with a number of additional challenges: currently, the input to the WFMM executable must be provided in the MATLAB (The MathWorks Inc., 2012) file format. This requirement may pose an additional barrier for users who are not familiar with or do not have access to MATLAB, since it is a proprietary software. While the WFMM outputs a number of summaries, these are returned as MATLAB files too. Further, users who wish to perform some custom inference on the posteriors will have to do extensive scripting in the programming language of their choice to access, organise and manipulate the raw posteriors.

Limitations in computational resources may become an issue when analysing data sets with high-dimensional curves and/or a large number of trials. The computational intensiveness of the WFMM can become a challenge and may put a limit on the size of the data and the model



complexity that can be accommodated in practice (see Herrick & Morris, 2006, for background information on computational considerations). Depending on the model complexity, WFMMs may be run on moderate-sized data sets (on the order of 10s to 100s of MBs) using standard laptops, but larger data sets may exceed available memory and may require other computing strategies. These include wavelet compression, grid computing facilities available with the WFMM software or sequential computational approaches. Also, complex models with large numbers of random effects may run more slowly and may take time to finish, since the computing time scales linearly with  $T$  (the number of data points per curve), but does not scale as well with the number of predictors or the number of random effects.

Further, the WFMM version used in the present work lacks model selection and evaluation procedures routinely applied with standard LMMs and does not yet offer the wide range of correlation structures available with many LMM packages. The current WFMM implementation is limited to independent random effects and residual errors, although Zhang et al. (2015) introduce a variant implementing functional conditional autoregressive (CAR) model for spatially correlated functional data and Zhu et al. (accepted) in very recent work present an FMM-based analysis of ERP-data taking spatial correlation between electrodes into account via Matérn assumptions in the basis space.

While the version of the WFMM software presented here is a very efficient and flexible tool for functional data analysis, users may thus face a number of limitations and challenges. For some of these, solutions will eventually be available, such as the approaches to modelling spatially correlated data mentioned above (Jeffrey S. Morris, personal communication). For potential new users of the software with little experience with LMMs and/or scripting in a suitable programming language, the biggest handicap is the lack of an intuitive user-interface which takes care of assembling the input to the WFMM executable and of handling its output in a flexible manner. The following two chapters present an overview two packages for the R statistical environment which act as such a WFMM wrapper for the analysis of general curve- and ERP-data.



## Chapter 7

# The *wrapfmm* package

### 7.1 Aims and scope

The aim of the *wrapfmm* package is to facilitate handling of Morris and Carroll's (2006) WFMM core software and to make it more accessible to the 'average' data analyst. In fact, one of the goals of the package is to make a WFMM analysis as intuitive as possible for an R user with working knowledge of the *lme4* (Bates et al., 2014) package for linear mixed modelling of scalar data. To this end, the package offers a number of functions which jointly act as a comprehensive user-friendly R wrapper for the WFMM core software. *wrapfmm* also considerably eases access to the WFMM by doing away with the requirement to have a valid MATLAB installation – writing and reading of MATLAB input and output files is handled internally by the *R.matlab* package (Bengtsson & Riedy, 2011).

While *wrapfmm* aims at making access to the core WFMM software as barrier-free as possible, it offers numerous tools which add to the functionality of the WFMM executable and allow a very flexible handling of the WFMM output. The focus here is on the *visual* display and exploration of the underlying functional data, a WFMM's posteriors and the joint plotting of data and model estimates. Another benefit of using *wrapfmm* is the flexibility it offers in computing custom post-hoc posteriors from the original fixed effect posteriors returned by the WFMM and in summarising and further exploring these. Further, the package comprises functions for basic model- and MCMC-diagnostics.

Key features of *wrapfmm* include:

- Specifying a WFMM via an *lme4*-style mixed model formula via the `wfmm` function.
- Full automation of the R/WFMM input/output pipeline, providing detailed status messages and an MCMC progress bar.
- Automatic computation of a number of default inferences such as pointwise credible intervals, simultaneous credible bands (SCBs) and different types of posterior prob-

abilities, including novel multiplicity-adjusted probabilities derived from the SCBs (see section 7.7).

- `plot` and `summary` methods for visualising the functional data and the WFMM estimates as high-quality plots using the *ggplot2* package (Wickham, 2009).
- Providing a high degree of flexibility for the user at all stages of analysis while minimising the amount of coding and previous knowledge required.

## 7.2 Availability & compatibility

As pointed out in the preface, the asynchrony of the package versions with current versions of R limits their practical value as presented here. However, they can be read as (extended) supplementary material to the ERP-analyses presented in section 4.2.6 of the first part and as complementary background chapters on new packages that will be available soon, retaining almost all of the basic functionality presented here, while at the same time offering many more features; for more information, including the status of the packages as presented here, please refer to the preface and the introductory comments to chapter 6.

## 7.3 Implementation

The *wrapfmm* interface is written in pure R code and currently does not make any ‘direct contact’ with the WFMM core executable. All exchange of information takes place making use of the existing WFMM architecture: the input to the WFMM executable is provided as a Matlab file and the relevant output is read into R from the WFMM files within the model directory. Status and error messages printed by the `wfmm` function are obtained by scanning the respective log and error files found in a model directory. This strict separation of the R layer and the core WFMM executable results in a certain degree of redundancy, by doubling a number of input/output structures within R and on disk. The *wrapfmm* objects, functions and classes are implemented within R’s *S3* system (e.g. Matloff, 2011, chapter 9).

## 7.4 WFMM backend

The WFMM backend has to be downloaded separately and is currently available at <https://biostatistics.mdanderson.org/SoftwareDownload>. Please note that the package versions presented here interface with the WFMM version 3.0. This version is only available for Windows.

## 7.5 Limitations

The *wrapfmm* package does not provide a complete interface to all WFMM variants, nor to all of the output returned by a call to the WFMM. While the latest version of the WFMM executable allows modelling of two-dimensional images and three-dimensional volumes, the *wrapfmm* functions only provide support for WFMMs of one-dimensional curve data. Note that among the additional functionality of the packages currently developed will be the possibility to fit multi-dimensional functional mixed models (see also the comments in the preface).

When called, the WFMM executable returns two MATLAB files containing a WFMM's input/starting parameters and summaries of its output, plus a number of 'flat' data files containing the posteriors of the fixed effects, variance components and random effect functions. All of these raw posteriors are available in the wavelet space, but only the posterior fixed effect curves are additionally provided in the original data space. Currently, *wrapfmm*'s `wfmm` method only reads in the data-space fixed effects posteriors and computes all default summaries (e.g. credible intervals and posterior probabilities) from these in R – at the moment, none of the *wrapfmm* functions provides access to any of the wavelet-space posteriors returned by the WFMM method, however.

Note that the only elements of the MATLAB files returned by the WFMM executable which are currently directly accessed by *wrapfmm* are the *rho* matrices and *sigma* vectors representing summary information about the model's variance components (see Morris & Herrick, 2012, for more detailed information) as well as the posterior mean estimates for the random effects curves (*U\_mean*), provided these were selected to be computed in the original call to `wfmm`. None of the other elements of the input/output MATLAB files is currently accessed by *wrapfmm*, but they can be easily read into R using the `readMat` function of the *R.matlab* package (Bengtsson & Riedy, 2011). The raw wavelet-space posteriors not accessed by *wrapfmm* can be read into R using the `readBin` function, but further processing will require custom scripts.

For more detailed information about the structure and content of WFMM input/output files as well as a complete list of WFMM settings, please refer to the WFMM user guide of Morris and Herrick (2012). For a discussion of inherent and practical limitations of the version of the WFMM core software used, refer to section 6.6.5.

Finally, please note that the WFMM software assumes that all curves contained in the data matrix  $Y$  are registered on a common grid  $t$ , i.e. that the observed functions already are on a common domain, with the single  $t$ 's (time points) corresponding across functions. If this is not the case, make sure that the observed curves are registered to a common grid in a preprocessing step using functions of your choice, since such a registration is not done by the WFMM software!

## 7.6 Overview

### 7.6.1 Getting oriented

The following is a brief list of the most relevant functions the package comprises. They are arranged in the 'order of appearance', reflecting the sequence of application within a typical workflow.

`wfmm_base` sets up a `wfmmBase` object which holds the matrix with the curve data, a data frame with variables reflecting related by-curve information as well as associated information about the data.

`plot` allows to plot curve data, providing an easy-to-use interface for breaking up the data by the levels of multiple factors, including on-the-fly conversion of numerical variables to factors and computation of difference curves. It is a generic function with methods for objects of class `wfmmBase`, `wfmm` and `postWfmm`. When applied to a `wfmm` or `postWfmm` object, the data plots can be combined with visualisations of posterior estimates of fixed effect curves. Its versatility makes this function a useful tool for exploring curve data prior to analysis and visualising a WFMM's output together with the underlying data.

`wfmm` is *wrapfmm*'s core function and is used for specifying and running a WFMM. It is applied to a `wfmmBase` object, the corresponding WFMM's structure is easily defined via an *lme4* mixed model statement (Bates et al., 2014) and options for the core WFMM executable can be specified conveniently. The function takes care of writing the required input MATLAB files with the data and model matrices via the *R.matlab* package (Bengtsson & Riedy, 2011), imports the model's posteriors to R, generates default inferences and returns an object of class `wfmm`.

`summary` returns a plot summary of a WFMM's posterior fixed or random effects curves or its associated variance components. For fixed effects, `summary` may also return a data frame with information for each significant cluster. Inference about fixed effects can be based on pointwise or simultaneous credible intervals or probabilities (see section 7.7 below for details). The `summary` function is a generic function with methods for `wfmm` and `postWfmm` objects, and thus also allows to summarise post-hoc posteriors computed with the `follow_up` function. It can further be used to return a summary plot of a WFMM's variance components.

`follow_up` is used to compute post-hoc posteriors from a WFMM's fixed effect posteriors, such as simple effects. The function returns an object of class `postWfmm` and all of the functions applicable to `wfmm` objects have appropriate methods for `postWfmm` objects. The `follow_up` function can also be used to compute 'higher order' post-hoc posteriors from post-hoc posteriors contained in another `postWfmm` object.

`chain_plot` is a generic function for making (informal) diagnostic plots for MCMC chains of fixed effect posteriors of a `wfmm` or `postWfmm` object. Plot types comprise chain plots, density plots and boxplots.

## 7.6.2 A quick tour

The current section will give a brief tour of *wrapfmm*'s core functions, providing code examples and pointing out relevant features. The data set used for illustration is the example data coming with the *wrapfmm* package, which is a small subset of a more comprehensive ERP experiment on semantic predictability of target words during visual sentence processing experiment, aiming at the N400 component. While this data set only contains factorial variables and no continuous numeric predictors, it is used here for a number of practical reasons, including ease of accessibility for users without access to the ERP data presented in this thesis. Also, most of the respective help pages provide further examples using the data, which makes it easy for users to directly expand upon the examples presented in the following.

### 7.6.2.1 The data set

The example data set used for demonstration here stems from an EEG sentence processing experiment (see Martín-Loeches et al., 2017)<sup>1</sup>, which comprised a total of 144 sentence units. The experiment aimed at the N400 component and predictability of the target words was manipulated such that they were either highly predictable or rather unpredictable given prior context, thus belonging either to the 'high' or 'low' predictability condition (see also Dambacher, Rolfs, Göllner, Kliegl, & Jacobs, 2009). The sentences containing the critical words were presented in rapid serial visual presentation. In the experiment, the 144 sentence units were split into two lists with 72 sentences each (containing 36 low- and high-predictability target words) and each list was presented to the same set of subjects in one of two sessions. The design was thus a 2x2 ('low' vs. 'high' predictability and 'session 1' vs. 'session 2') factorial design.

During the experiment, the ongoing EEG was recorded at 27 electrode sites. The 1100 ms ERP segments used for analysis were cut around the critical words with a pre-stimulus baseline of 200 ms, thus extending 900 ms post stimulus onset. Segments were baseline-corrected against the 200 ms pre-stimulus baseline period and represent average-referenced ERPs sampled at 250 Hz. The example data set used in the following comprises a small subset of single-trial data from eight subjects and one posterior midline electrode, Pz. Overall, the data set comprises 1063 ERP segments ( $n = 1063$ ) with 275 data points each (i.e.  $T = 275$ ).

---

<sup>1</sup>I would like to thank Professor Werner Sommer for making part of the data available for the current purposes.

### 7.6.2.2 wfmm\_base: setting up a base object

The first step in using *wrapfmm* is setting up a 'base' object via the *wfmm\_base* function. The resulting object holds the matrix with the curve data (Y), a data frame with corresponding variables providing information about each curve as well as information about the dimensionality of the data and the associated units. The base object is constructed with the following call – once available, the *wfmm* and *plot* methods can be applied to a base object, to run a WFMM and plot the underlying curve data, respectively.

```
n400.pz.base <- wfmm_base(Y = n400.pz.data,
                        variables = n400.pz.vars,
                        min.t = -200,
                        max.t = 896,
                        units = list(x = "ms",
                                    y = expression(paste(mu, "V"))))
```

### 7.6.2.3 wfmm: running a wavelet-based functional mixed model

The *wfmm* function is used to actually run a WFMM – it automatically calls the WFMM core executable, manages the respective input/output files on disk, imports a number of summary statistics returned by the WFMM, imports the raw fixed effects posteriors, computes a number of default inferences from them (see section 7.7 below) and returns an object of class *wfmm*, for which numerous methods are available (see section 7.6.1 above). A WFMM is defined using the mixed model formula interface applied in the *lme4* package (Bates et al., 2014), which allows easy and flexible specification of the fixed and random effects parts of a mixed effects model.

The following call defines a WFMM with an intercept effect plus effects for the 'c.Session' and 'c.Predictability' variables as well as their interaction in the fixed effects part. Note that the 'c.' prefix indicates that these are transformed versions of the original 'Session' and 'Predictability' variables which have been centred on their mean. On the random effects side, the model features by-subject random intercepts and by-subject random slopes for the 'c.Predictability' effect. Note that the *disk.id* argument provides the name used for the model directory and the related WFMM files on disk. While it does not have to match the name of the object returned by *wfmm*, using the same name is recommended to ensure naming consistency.

```
n400.pz.wfmm <-
  wfmm(x = n400.pz.base,
      formula = ~ 1 + c.Session * c.Predictability +
        (1 + c.Predictability | Subject),
      disk.id = "n400.pz.wfmm",
      PostProcessSpecs = list(compute_U = 1),
      MCMCSpecs = list(B = 2000, burnin = 1000, thin = 3),
      wavespecs = list(wavelet = "sym8"))
```



While `wfmm` makes use of `lmer`'s LMM formula interface, there is one crucial difference residing in the specification of the random effects part which users should be aware of: when using `lmer`, a random effects specification such as  $(1 + c.Predictability \mid Subject)$  fits a model with correlated random effects, since the by-subject random intercept and slope are included within the same set of parentheses. Since the WFMM currently can only fit models with independent random effects, the above specification is actually only shorthand for the variant which explicitly fits a model *without* correlated random effects, i.e.  $(1 \mid Subject) + (0 + c.Predictability \mid Subject)$  for the above case.

The `wfmm` call also illustrates a feature of the *wrapfmm* user interface: many of its functions take lists with named list entries of various types as arguments. Sometimes, such list arguments may have numerous entries defining options or other parameters which have predefined default settings. For these, users can only change a subset of these parameters in a given call and any of the non-adapted list entries will be merged in with their default settings. The `PostProcessSpecs`, `MCMCSpecs` and `wavespecs` list-arguments illustrate these cases: all of them may contain more entries than specified and in all cases the non-specified entries are used internally with their default values. This way, *wrapfmm* tries to strike a balance between offering a very high degree of flexibility/customisability and retaining a compact user interface and keeping the amount of required code manageable.

#### 7.6.2.4 `summary`: summarising a WFMM's posterior estimates

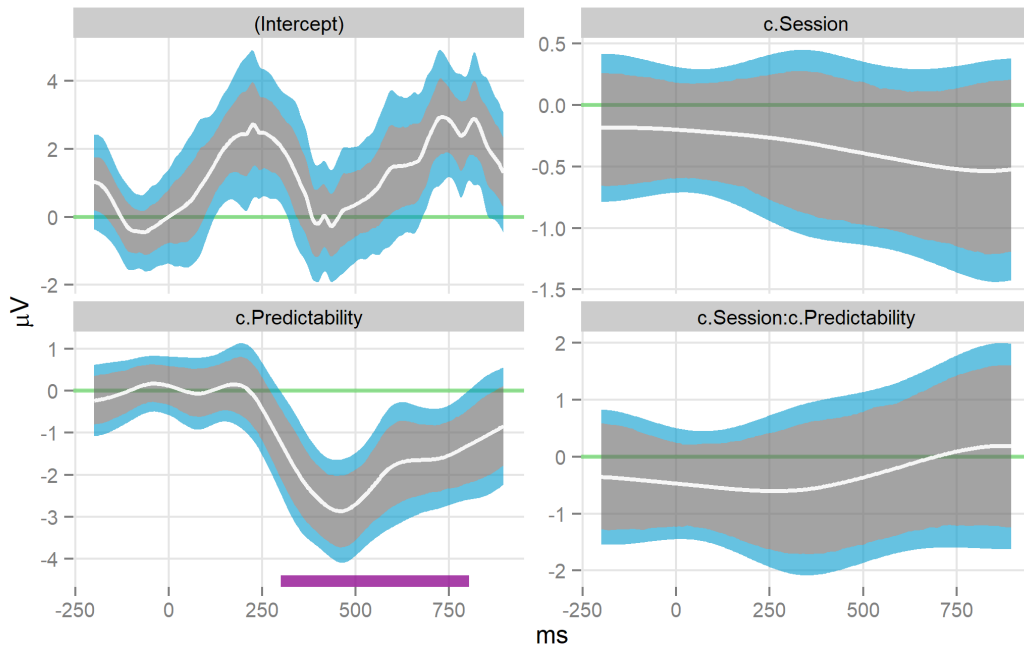
The `summary` method allows to summarise a number of a WFMM's posterior estimates, including fixed and random effects as well as variance component estimates. The richest set of options for summary is available for fixed effects, including a number of different inference bases such as pointwise CIs, SCBs as well as associated posterior probabilities. Summaries of the random effects and variance components are currently limited to the summary statistics returned by the core WFMM executable. Detailed information about the computation and interpretation of the different inference bases available for fixed effects can be found in section 7.7 – here, I will only provide a superficial description of their properties. In contrast to most other `summary` methods available in other packages, the default output of *wrapfmm*'s `summary` is a plot, rather than a summary in text or table format, to do justice to the complex functional nature of the data and model.

```
summary(n400.pz.wfmm)
```

The above call returns the plot rendered as figure 7.1, which summarises all four fixed effects of the WFMM. By default, inference about significance of fixed effects is based upon *type II* SCBs (see section 7.7.1.2), which are plotted in blue (grey bands indicate pointwise CIs) – where these do not contain zero, a bar at the bottom of each panel marks the respective significant intervals<sup>2</sup>. In the case at hand, the only significant cluster belongs to the `c.Predictability` effect: the lilac bar marks a negative-slope effect starting at about 300 ms and lasting until about 800 ms with a peak at around 450 ms.

<sup>2</sup>Note that for (Intercept) fixed effects, no significance indicator bars are added to the respective panel.

## Fixed effects posterior summary for 'n400.pz.wfmm'

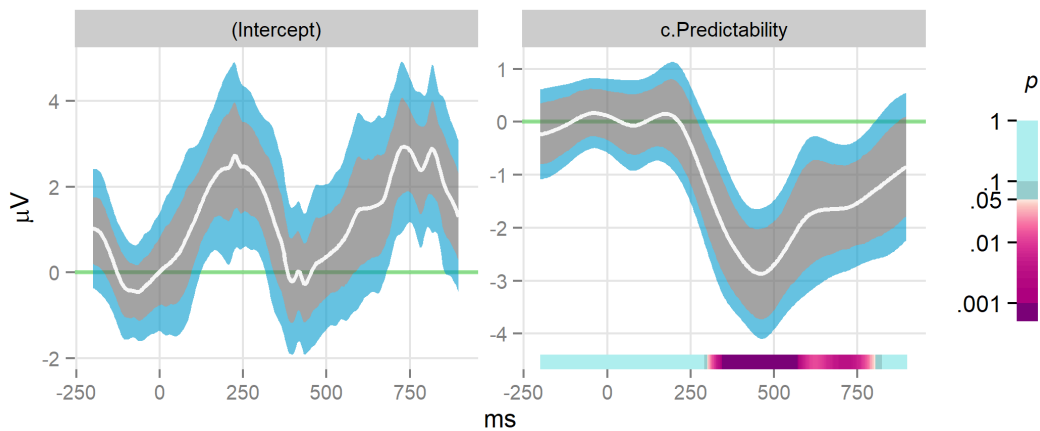


**Figure 7.1:** Posterior fixed effects summary plot of the example WFMM with default settings. Assessment of significance is based on whether the type II SCBs (outer blue bands) contain zero or not.

Each of the CIs/SCBs available for summarising fixed effects comes with a corresponding posterior probability. These allow to base inference about significance on continuous numerical probabilities which aim at flagging the same set of time points as significant as their corresponding CI/SCB (see section 7.7.2 for details). For the following summary, the inference basis is set to the probabilities derived from the *type II* SCBs, the 'simultaneous band scores' (SIMBAS) of Meyer et al. (2015). By default, the threshold  $\alpha$  for significance is set to the complement of the CI's/SCB's probability, i.e.  $\alpha = 1 - 0.95 = 0.05$  for the current case. Figure 7.2 shows the output of the following call, which selects the (Intercept) and c.Predictability predictors for summary.

```
summary(n400.pz.wfmm,
  ffx.settings = list(predictors = c("(Intercept)",
                                    "c.Predictability"),
    inference.basis = "prob"),
  title = NULL)
```

While the primary mode for WFMM summaries is graphical, for some data sets summary plots with default settings may be of limited value – for high-dimensional spiky and irregular curves with thousands of sampling points, summary plots covering the complete grid may be



**Figure 7.2:** Posterior fixed effects summary plot for the (Intercept) and Predictability effects with inference based upon SIMBAS.

hard to parse and short significant clusters may be missed. As an alternative, cluster-level summary statistics in form of a data frame for any of the inference bases available with the `summary` method can be obtained by setting the output argument to "table". Subsequently, any intervals of interest may be summarised visually by zooming into the x-axis using the `x.zoom` argument. The following call returns the summary data frame presented below.

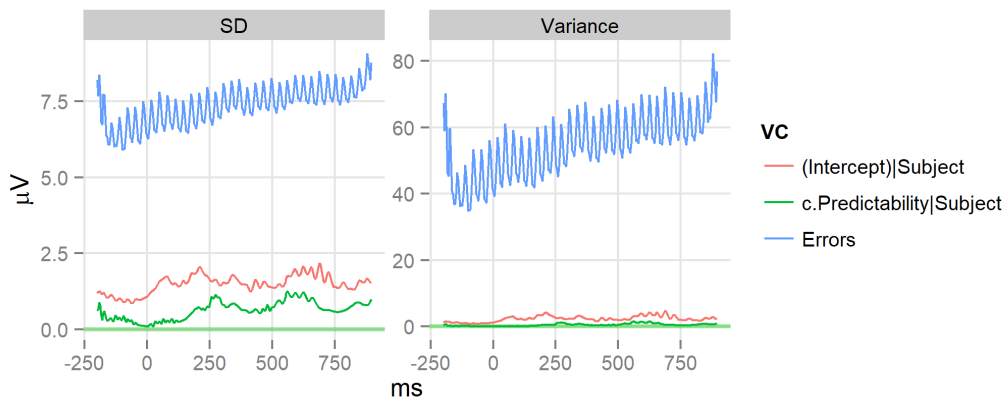
```
summary(n400.pz.wfmm,
        output = "table",
        ffx.settings = list(inference.basis = "prob"))
```

	Fixed FX	Start	End	Length	Mean est.	min.	Mean est.	max.
1	c.Predictability	300	804	504	-2.87		-1.24	
		Est. max @ P.						
1		460	0.00025					

Finally, using the `type` argument, `summary` can summarise posterior estimates for a WFMM's variance components (figure 7.3) as well as random effects (provided that the `compute_U` switch of the `PostProcessSpecs` argument was actually set to 1 in the original call to `wfmm` – see figure 7.4):

```
# Variance component summary
summary(n400.pz.wfmm,
        type = "VC")

# Random effects summary
summary(n400.pz.wfmm,
        type = "random")
```



**Figure 7.3:** Summary of posterior mean estimates of the WFMM's variance components.

*wrapfmm*'s `summary` method is one of the package's core utilities, providing an easy-to-use yet flexible function for the most important estimates of a WFMM. While it aims at using sensible default settings for the different summary types, users have extensive control over a large number of settings, including subsetting, panel layouting via the 'facetting' interface of the *ggplot2* package (Wickham, 2009) and numerous graphics options.

#### 7.6.2.5 `follow_up`: computing post-hoc posteriors

Once a WFMM has been fitted, the `follow_up` function is applied to compute custom post-hoc posteriors from a WFMM's fixed effects posteriors. In the following, we compute the cell-mean estimates for 'low' and 'high' predictability words from model `n400.pz.wfmm`. Since the model only features mean-centred predictors, the (Intercept) estimate does not indicate a specific reference level, but rather the overall mean curve. Thus, the cell means estimates are obtained by adding/subtracting the appropriately weighted posteriors of the `c.Predictability` predictor to the (Intercept) posteriors. For the specification of follow-up posteriors, two functions can be used: `beta`, which serves to access a given fixed effect's raw posterior samples and `pred`, which accesses the respective column of the fixed effects design matrix  $X$ . Prior to the call to `follow_up`, the *plyr* package is loaded to make the `.`() function available, which simplifies the specification of the `formulas` argument.

```
require(plyr)

n400.pz.postwfmm <-
  follow_up(n400.pz.wfmm,
    disk.id = "n400.pz.postwfmm",
    formulas = list(High = .(beta((Intercept)) +
      beta(c.Predictability) * min(pred(c.Predictability))),
      Low = .(beta((Intercept)) +
      beta(c.Predictability) * max(pred(c.Predictability)))))
```

`follow_up` returns an object of class `postWfmm`, for which the same range of methods as

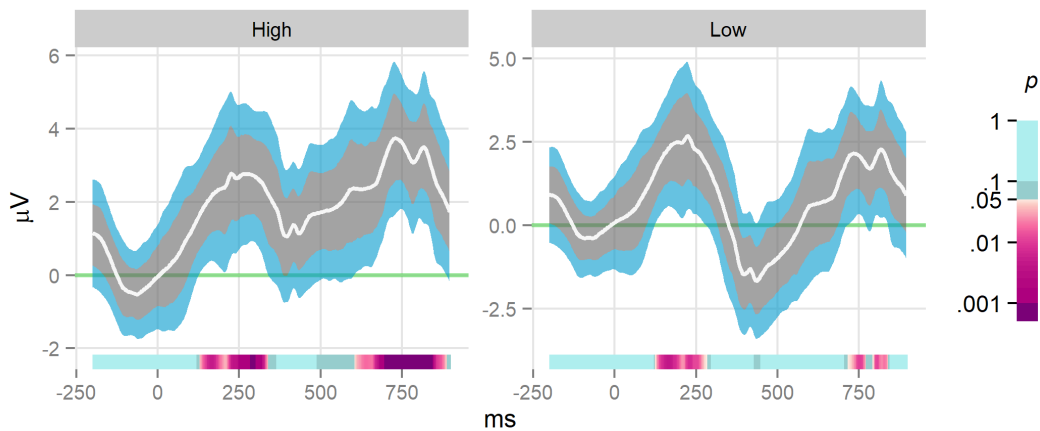


**Figure 7.4:** Summary of the example WFMM's posterior by-subject random effects estimates (mean and pointwise CIs for the random intercepts and random slopes for the Predictability effect).

for `wfmm` objects is available. The `summary` method can be applied to `postWfmm` objects to produce summaries with the same range of inference bases and options available for fixed effects summaries of `wfmm` objects. The below call to `summary` produces the summary plot presented in figure 7.5:

```
summary(n400.pz.postwfmm,
       ffx.settings = list(inference.basis = "prob"))
```

Also, 'higher-level' follow-up posteriors can be computed from previously computed post-hoc posteriors by either applying `follow_up` to a `postWfmm` object, or by using previously defined follow-up posteriors 'on the fly' within the same function call. Thus, `follow_up` is a key tool for further mining a WFMM's posteriors in a highly flexible manner. While common applications include computing lower-level effects in the presence of higher-order interactions, its interface is general enough to allow creative specification of a wide variety of follow-up posteriors.



**Figure 7.5:** Summary plot of the posterior cell means estimates for low and high predictability words computed with the `follow_up` function (inference based upon SIMBAS).

#### 7.6.2.6 `plot`: plotting data and/or fixed effects posteriors

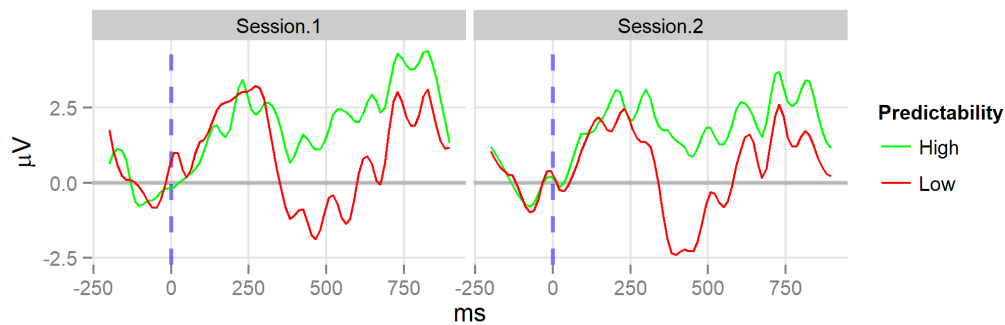
The `plot` function provides methods for a number of different classes of the `wfmm` family: when applied to a `wfmmBase` object, it allows to explore the respective curve data prior to data analysis in a very flexible manner. When used with objects of class `wfmm` or `postWfmm`, `plot` can be applied to graph curve data together with the respective posterior fixed effects estimates. The function's user interface again provides a high degree of flexibility, rendering it a versatile tool for data and model exploration.

The first example call applies the function to the `wfmm` base and plots the mean curves for each level of the 'Predictability' factor, mapping them onto different colours. In addition, the `facet.wrap` formula sets up two separate panels for each level of the 'Session' factor. The resulting figure 7.6 thus shows the mean curves for 'Predictability' by 'Session'.

```
plot(n400.pz.base,
     facet.wrap = ~ Session,
     data.groups = list(Predictability =
                        list(aes = "colour",
                             mapping = c("Green", "Red"))),
     v.line = list(list(intercept = 0, linetype = 2)),
     smooth.curves = list(smooth = TRUE),
     free.y = FALSE)
```

Using the `diffcurves` argument, difference curves can be computed easily and added to the plot (figure 7.7):

```
plot(n400.pz.base,
     diffcurves = list("Low-High" = list(
```

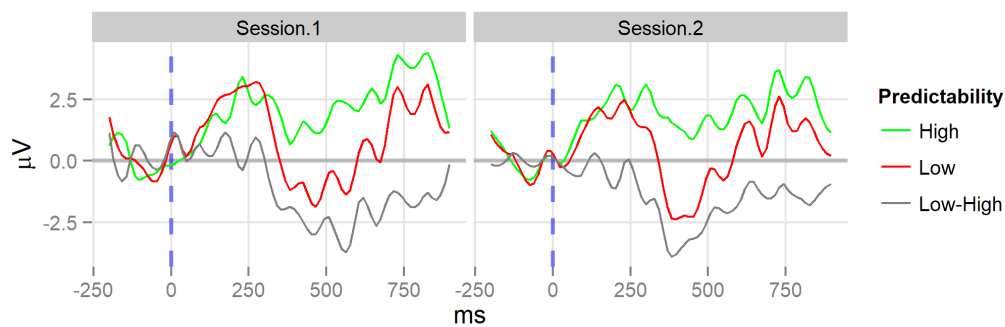


**Figure 7.6:** Mean curves for the 'Low' and 'High' levels of the 'Predictability' factor, faceted by 'Session'.

```

factor = "Predictability",
difference = c("Low", "High"))),
facet.wrap = ~ Session,
data.groups = list(Predictability = list(
  aes = "colour",
  mapping = c("Green", "Red", "Grey50"))),
v.line = list(list(intercept = 0, linetype = 2)),
smooth.curves = list(smooth = TRUE),
free.y = FALSE)

```



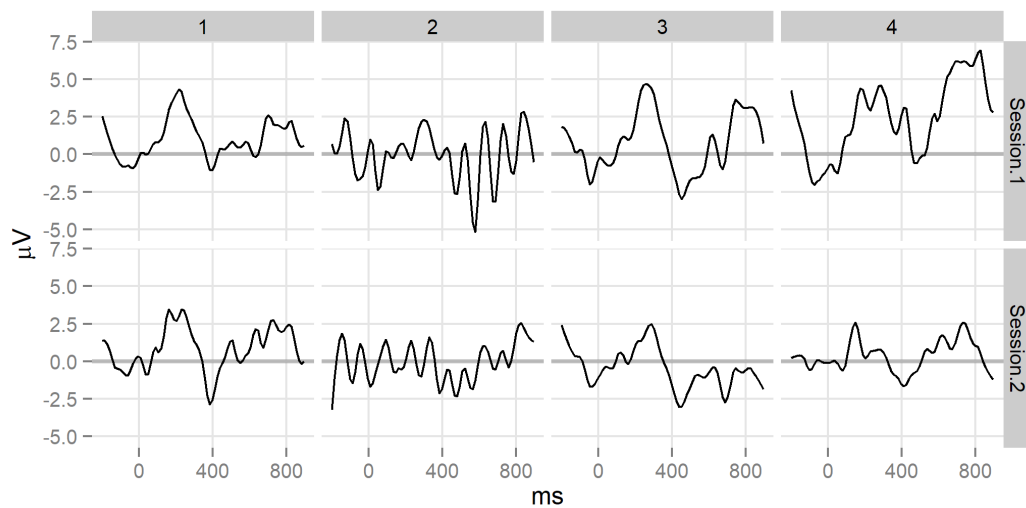
**Figure 7.7:** Same as in previous figure, with difference curves added.

Note that facetting formulas may contain multiple factors – in the following example call, the difference curves for the 'Predictability' factor are plotted separately for each of a subset of four subjects using `facet.grid` to inspect inter-individual variability within and across sessions (see figure 7.8). In fact, any combination of variables encoded as factors can be used for mapping of the respective levels and continuous numerical variables can be converted to factors on the fly.

```

plot(n400.pz.base,
     facet.grid = Session ~ Subject,
     subset = .(Subject %in% 1:4),
     diffcurves =
       list("Low-High" = list(
         factor = "Predictability",
         difference = c("Low", "High"))),
     smooth.curves = list(smooth = TRUE),
     x.breaks = c(0, 400, 800),
     free.y = FALSE)

```



**Figure 7.8:** Difference curves for the two levels of the 'Predictability' factor faceted by 'Session' and a subset of four subjects.

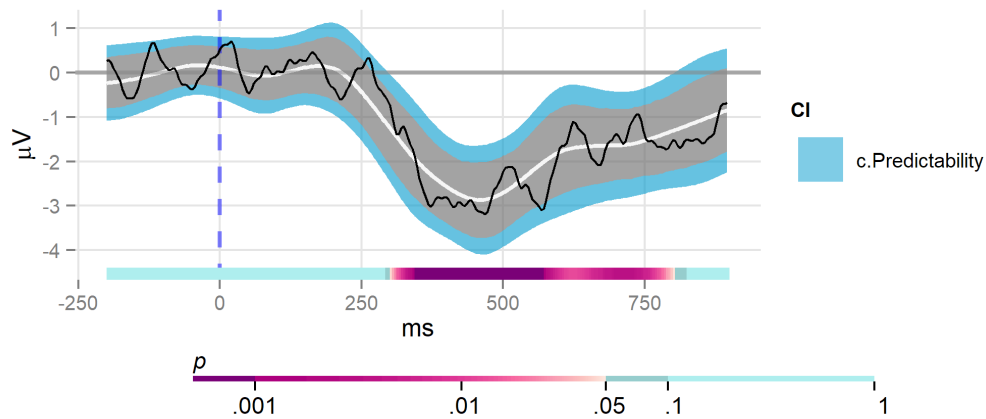
In the below call, `plot` is applied to the `wfmm` object returned by the original call to the `wfmm` function. The resulting plot (figure 7.9) compares the overall difference curves for the 'Predictability' factor with the estimate for the `c.Predictability` predictor, again selecting SIMBAS as inference basis.

```

plot(n400.pz.wfmm,
     ci = list('c.Predictability' = list()),
     ffx.settings = list(inference.basis = "prob"),
     diffcurves =
       list("Low-High" = list(
         factor = "Predictability", difference = c("Low", "High")),
         plot.diffcurves.only = TRUE),
     v.line = list(list(intercept = 0, linetype = 2)))

```



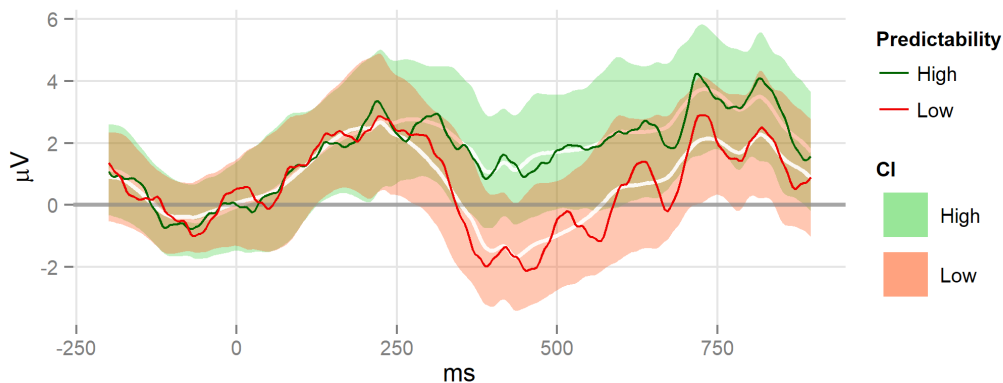


**Figure 7.9:** Difference curve for 'High' and 'Low' predictability words compared to the posterior estimate for the Predictability fixed effect.

The last plot example (figure 7.10) illustrates how to overlay data-based curves with posterior estimates contained in a `postwfmm` object: here, the cell mean curves for the 'High' and 'Low' levels of the 'Predictability' factor are compared to the respective cell means estimates computed by the call to `follow_up`, as described above, using `type // SCBs` for the estimates.

```
plot(n400.pz.postwfmm,
     ci =
       list('High' = list(colour = c("limegreen"),
                           mean.col = "white",
                           alpha = 0.3),
            'Low' = list(colour = c("orangered"),
                          mean.col = "white",
                          alpha = 0.3)),
     data.groups =
       list(Predictability = list(
         aes = "colour",
         mapping = c("darkgreen", "red2"))),
     ffx.settings = list(inference.basis = NULL),
     ci.type = "sim.II")
```

Thus, the plot methods complement the other functions presented at any step of analysis: they are valuable tools for exploring the curve data prior to fitting a WFMM or for comparing data and model estimates once a WFMM has been run.



**Figure 7.10:** Cell mean curves for the two levels of the 'Low' and 'High' levels of the 'Predictability' factor together with the corresponding posterior follow-up estimates.

## 7.7 Posterior inference: computation of credible intervals and posterior probabilities

### 7.7.1 Credible intervals

#### 7.7.1.1 Pointwise credible intervals

The `wfmm` function automatically computes pointwise credible intervals and simultaneous credible bands ( $pwCI$  and  $SCB$ , respectively) for posterior fixed effect curves from the raw MCMC samples of the fixed effects. The  $pwCI$ s are quantile-based credible intervals at a defined target probability content, which is defined using the `prob` argument in the call to `wfmm` or `recap`. Note that the same value will be used for the computation of the simultaneous credible bands, i.e.  $prob_{pwCI} = prob_{SCB}$ . If not mentioned otherwise, I will make use of 95% pointwise credible intervals and SCBs ( $0.95 = prob_{pwCI} = prob_{SCB}$ ) in all subsequent examples, which is the default setting.

#### 7.7.1.2 Simultaneous credible bands

`wrapfmm` offers two different SCBs for inference about fixed effect curves, both of which take into account multiplicity along the grid  $\mathbf{t}$ . The *type I* simultaneous Bayesian credible bands are derived using a rather intuitive method of computation presented in Krivobokova et al. (2010, section 3): in a first step, the algorithm starts with computing pointwise lower and upper quantiles from the sampled curves, which are then iteratively scaled by a constant factor. The algorithm stops when at least  $prob_{SCB}100\%$  of all sampled curves are fully contained within the resulting bands.

This method was first applied to posterior fixed effect curves of a WFMM by Steen (2010). As pointed out by Krivobokova et al. (2010), their approach to computation of simultan-

eous Bayesian credible bands is related to the one described by Ruppert, Wand, and Carroll (2003) and Crainiceanu, Ruppert, Carroll, Joshi, and Goodner (2007). However, since the latter is based on the pointwise posterior means and standard deviations, it has the drawback of implicitly relying on posterior normality, an assumption which is avoided in Krivobokova et al.'s method<sup>3</sup>.

In addition to *type I* simultaneous credible bands, *wrapfmm* also implements Ruppert et al.'s variant as *type II* simultaneous credible bands (Ruppert et al., 2003). Using the notation of Meyer et al. (2015), who construct these simultaneous credible bands for WFMM-based function-on-function regression, these are calculated as

$$I_\alpha(t_j) = \hat{\beta}_i(t_j) \pm q_{(1-\alpha)} \left[ \widehat{St.Dev} \left\{ \hat{\beta}_i(t_j) \right\} \right], \quad (7.1)$$

where  $\hat{\beta}_i(t_j)$  and  $\widehat{St.Dev} \left\{ \hat{\beta}_i(t_j) \right\}$  are the mean and standard deviation<sup>4</sup> of all  $G$  MCMC samples at  $t_j$  for fixed effect  $\beta_i$  and the variable  $q_{(1-\alpha)}$  is calculated as the  $(1 - \alpha)$  quantile of

$$Z^{(g)} = \max_{t \in T} \left| \frac{\beta_i^{(g)}(t_j) - \hat{\beta}_i(t_j)}{\widehat{St.Dev} \left\{ \hat{\beta}_i(t_j) \right\}} \right|. \quad (7.2)$$

### 7.7.2 Posterior probabilities

For each of the three different types of credible intervals/bands presented, *wrapfmm* provides a related posterior probability type which can be chosen for inference about fixed effects. The aim of each probability type is to allow inference based on continuous numerical probabilities which flag the same set of grid points as significant as their corresponding CI/SCB-counterparts (or, at a minimum closely approximate the respective set of significant grid points), given that the significance threshold  $\alpha$  selected equals the complement of the CI's/SCB's target probability content, i.e.  $\alpha = 1 - prob_{pwCI/SCB}$ .

<sup>3</sup>As pointed out to me in a personal communication by Jeffrey S. Morris, however, the approach to SCB computation of Krivobokova et al. may not perform optimally with very spiky functions. In such cases, the shape of the pointwise credible bands may vary a lot with  $prob_{pwCI}$ , such that multiplying them by a constant may result in poor operating characteristics.

<sup>4</sup>As Jeffrey S. Morris (personal communication) pointed out to me, estimation of the SD may require a correction when there are a high degree of autocorrelation in the MCMC chain AND a low number of MCMC samples. In web appendix G of the supplementary materials, Meyer et al. (2015) present a formula to perform this correction in such situations, but the *wrapfmm* package does not provide a way of using these corrected SDs for computation of the SCBs.

### 7.7.2.1 Pointwise posterior probabilities

The pointwise posterior probabilities ( $pw.p$ ) for each grid point  $t_j$  of a given fixed effect  $\beta_i$  are computed as

$$pw.p_{\beta_i}(t_j) = 2 \cdot \min [\Pr \{\beta_i(t_j) > 0 \mid Y\}, \Pr \{\beta_i(t_j) < 0 \mid Y\}], \quad (7.3)$$

where  $G$  is the number of MCMC samples obtained and  $Y$  the data. Subsequently, any  $pw.p_{\beta_i}(t_j) = 0$  are replaced with  $(2 \cdot G)^{-1}$ . This is essentially the approach outlined in Fazio et al. (2014), who also point out that the resulting probabilities approximately correspond to a classical  $p$ -value for a two-sided hypothesis test of the null hypothesis that  $\beta_i(t_j) = 0$ , given non-informative priors for the fixed effects regression coefficients (see also Morris, Brown, et al., 2006).

### 7.7.2.2 ‘Simultaneous posterior probabilities’

For each of the *type I* and *type II* SCBs, *wrapfmm* computes corresponding ‘simultaneous probabilities’ ( $sim.p$ ) for posterior fixed effects estimates. These reflect posterior probabilities correcting for multiple comparisons across the grid  $t$  without the need to specify minimum effect sizes (see Morris et al., 2008, p. 484, for a method to calculate probabilities using a combination of a false discovery rate – FDR – and minimum effect size) .

The probabilities based upon the *type II* simultaneous credible bands introduced by Meyer et al. (2015) are termed SIMBAS (‘**S**imultaneous **B**and **S**cores’) and are computed by *wrapfmm*’s `wfmm` function as

$$SIMBAS_{\beta_i}(t_j) = \frac{1}{G} \sum_{g=1}^G 1 \left\{ \left| \frac{\hat{\beta}_i(t_j)}{\widehat{St.Dev} \{ \hat{\beta}_i(t_j) \}} \right| < Z^{(g)} \right\},$$

with the quantities  $\hat{\beta}_i(t_j)$ ,  $\widehat{St.Dev} \{ \hat{\beta}_i(t_j) \}$  and  $Z^{(g)}$  having the same interpretations as in equations 7.1 and 7.2 above. In the *wrapfmm* implementation, any  $SIMBAS_{\beta_i}(t_j) = 0$  are subsequently again replaced with  $(2 \cdot G)^{-1}$ .

In the following, I will present the method used for computing simultaneous probabilities derived from *type I* credible bands in the *wrapfmm* package – since this is an experimental approach which has not undergone external review yet, I will go into some detail to illustrate its properties and conceptual justification. The key idea for the computation of simultaneous probabilities as presented here is to set up an objectively determined ‘region of practical equivalence’ (or ROPE) along  $t$ . According to Kruschke (2010b, p. 301), a ROPE “indicates a small range of values that are considered practically equivalent to the null value for purposes of the particular application”. Thus, setting up a ROPE  $\delta$  amounts to defining an interval around zero (or any other parameter value of interest  $\theta$ ), with lower limit  $\delta_{Lower}^\theta$  and upper limit  $\delta_{Upper}^\theta$ , and treating any values within  $\delta^\theta$  as equivalent to  $\theta$  for any further inference about the credibility of  $\theta$ .

Kruschke (2010a, 2010b) presents a decision rule according to which  $\theta$  is NOT credible – i.e. rejected – if  $\theta$ 's entire ROPE lies outside of the credible interval of the respective parameter's posterior distribution. Note that the application of a ROPE also allows to declare  $\theta$  to be accepted (i.e. credible), which is taken to be the case if a ROPE *completely contains* a parameter's credible interval – thus, a null value  $\theta$  cannot only be rejected, but may also be *accepted*. As Kruschke (2010b) points out, such binary decisions about a parameter value fail to make use of the rich information provided by the posterior distribution. However, if a given parameter's credible interval overlaps with a ROPE, but is not completely contained by it, decisions about the credibility of the parameter value of interest may be suspended.

The width of any ROPE may be determined in a number of ways: for example, in clinical studies assessing the utility of a new drug, expert interviews may be conducted and used to reach a consensus about a minimal effect size which will justify switching to the new drug tested (taking into account cost/benefit considerations such as enhanced risk of side effects). In other settings, the decision about a ROPE's width may less be based on practical considerations and will tend to involve somewhat more arbitrary criteria. Either way, setting up a ROPE usually involves subjective decisions by the researcher(s). The key idea for the computation of simultaneous probabilities as presented here is to resort to the *type I* SCBs computed and use the differences between the lower and upper limits of the pointwise credible intervals and the respective counterparts of the SCBs for setting up an *objectively* determined functional version of a ROPE along  $\mathbf{t}$ , which I will refer to as 'BOND' (**B**and **O**f **N**o **D**ifference').

For the current purposes, I take the parameter value of interest to be zero at each  $\mathbf{t}$  for any functional fixed effect  $\beta_i$  (i.e.  $\theta_{\beta_i} = 0$ ); a BOND for any functional fixed effect  $\beta_i$  has the lower functional limit  $BOND_{\beta_i}^{Lower}$  and upper functional limit  $BOND_{\beta_i}^{Upper}$ , which are computed at any grid point  $t_j$  as stated in equations 7.4 and 7.5:

$$BOND_{\beta_i}^{Lower}(t_j) = -1 \cdot (SCB_{\beta_i}^{Upper}(t_j) - pwCI_{\beta_i}^{Upper}(t_j)) \quad (7.4)$$

$$BOND_{\beta_i}^{Upper}(t_j) = -1 \cdot (SCB_{\beta_i}^{Lower}(t_j) - pwCI_{\beta_i}^{Lower}(t_j)), \quad (7.5)$$

where  $pwCI_{\beta_i}^{Lower}(t_j)$  and  $pwCI_{\beta_i}^{Upper}(t_j)$  denote the lower and upper limits of the pointwise credible intervals of a functional fixed effect  $\beta_i$  at grid point  $t_j$ , respectively, and  $SCB_{\beta_i}^{Lower}(t_j)$  and  $SCB_{\beta_i}^{Upper}(t_j)$  the counterparts of the *type I* SCBs. The respective BOND is then defined as  $BOND_{\beta_i}(t_j) = [BOND_{\beta_i}^{Lower}(t_j), BOND_{\beta_i}^{Upper}(t_j)]$  and  $BOND_{\beta_i} = [BOND_{\beta_i}(t_1), \dots, BOND_{\beta_i}(t_T)]$ . Note that the resulting bands may vary in their width across  $\mathbf{t}$  and may not be fully symmetric, reflecting properties of the underlying posterior.

I illustrate the rationale for this first step in computing *type I* SCB-derived simultaneous probabilities by taking a look at the single significant effect cluster of the 'c.Linking' fixed effect of the WFMM as defined in section 4.2.6 within the first part of this thesis; while the original model was run for each of 63 EEG-electrodes, the following examples are based on the single WFMM for the anterior midline electrode AFz. Basing the assessment of significance on whether the *type I* SCB includes zero or not, the first line of table 7.1 provides

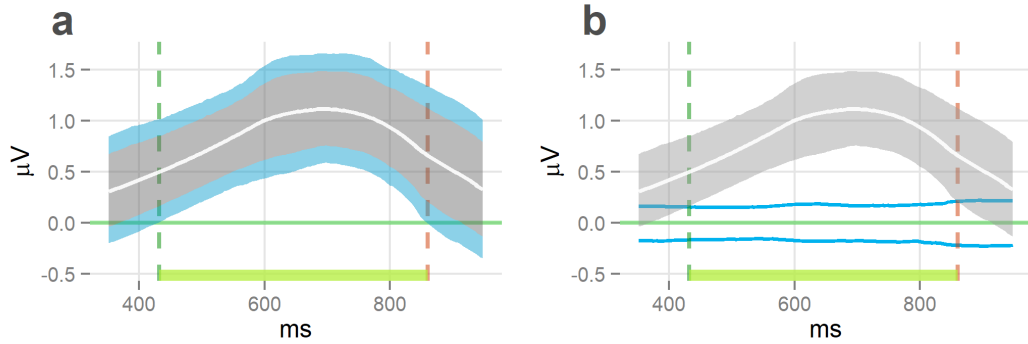
the corresponding start/end points of this cluster (432 ms/860 ms) as well as its length (428 ms). Figure 7.11a shows the respective pointwise (grey) and simultaneous (blue) credible intervals with the corresponding green bar indicating a positive-slope effect. The broken green/red lines mark those time points where the simultaneous credible band crosses zero, i.e. the start and end points of the effect cluster.

Figure 7.11b again shows the pointwise credible intervals in grey, while the blue lines below and above zero represent the lower and upper limits of the BOND. Here, the distance between zero and the *upper* BOND limit represents the difference between the *lower* simultaneous and *lower* pointwise credible interval limits, while the distance between zero and the *lower* BOND limit represents the difference between the respective *upper* credible interval limits. If assessment of significance is then based on whether the *pointwise* credible interval does NOT include the BOND, it is easy to see from figure 7.11b that this approach will yield conclusions which are equivalent to those based on the method underlying figure 7.11a: in fact, the two approaches outlined here and illustrated in figure 7.11 are just alternative (and equivalent) ways of accounting for the 'penalty' imposed by the multiplicity correction for SCB-based inference.

Fixed Effect	Inference basis	Cluster	Start	End	Length
c.Linkng	SCB I	1	432	860	428
	BONDS	1	432	860	428
	SCB II	1	440	860	420
	SIMBAS	1	440	860	420
c.Linkng:s.Al	SCB I	1	520	652	132
	BONDS	1	520	652	132
	SCB II	1	516	648	132
	SIMBAS	1	516	648	132
	SCB I	2	688	748	60
	BONDS	2	688	748	60
	SCB II	2	692	752	60
	SIMBAS	2	692	752	60
s.Word.Length	SCB I	1	364	524	160
	BONDS	1	364	524	160
	SCB II	1	360	528	168
	SIMBAS	1	360	528	168
	SCB I	2	560	772	212
	BONDS	2	560	772	212
	SCB II	2	560	776	216
	SIMBAS	2	560	776	216

**Table 7.1:** Details for the effect clusters for electrode FCz. Cluster information is based on four different inference bases: Whether the type I/II SCBs include zero, MULTIBONDS and SIMBAS.

Since  $BOND_{\beta_i}$  thus represents the multiplicity correction induced by the SCB of a functional posterior fixed effect  $\beta_i$ , a BOND can be used to compute *type I* SCB-based simultaneous probabilities, which I will call MULTIBONDS (**M**ultiplicity-Induced **B**and **O**f **N**o **D**ifference



**Figure 7.11:** a: Pointwise (grey bands) and *type I* SCBs (blue bands) for the 'c.Linking' predictor (350–950 ms). A time point is flagged as significant if the simultaneous credible band does not contain zero. b: Pointwise CIs and lower/upper limits of the functional BOND (blue lines). A time point is flagged as significant if the pointwise CI does not overlap with the BOND. In both plots, the broken green and red lines indicate the start/end points of the respective effect cluster and the green bar at the bottom marks this interval.

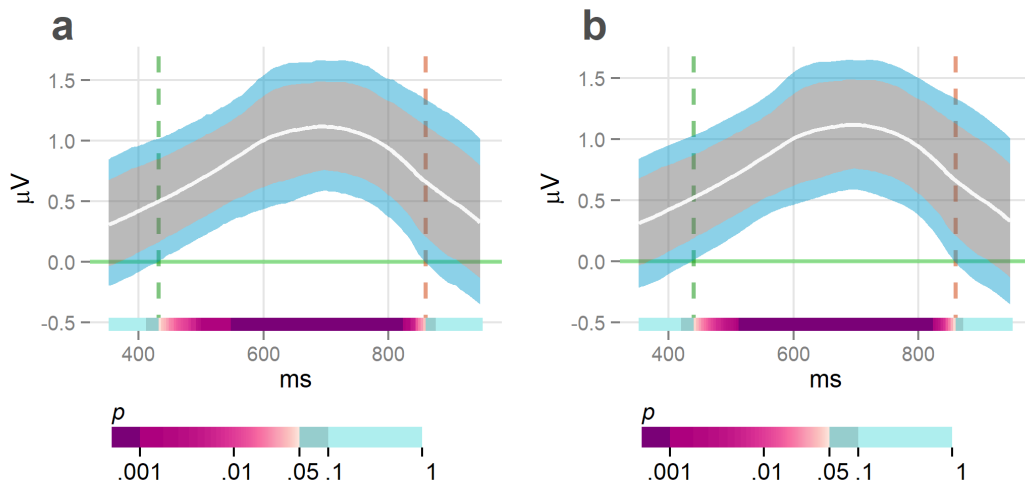
Scores) or BONDS, in short. MULTIBONDS reflect the probability of  $\beta_i(t_j)$  being equivalent to the parameter value of interest  $\theta$  given  $BOND_{\beta_i}(t_j)$  and are computed as

$$\begin{aligned} MULTIBONDS_{\beta_i}(t_j) &= \Pr \{ \beta_i(t_j) \in BOND_{\beta_i}(t_j) \mid Y \} \\ &\approx 2 \cdot \left( G^{-1} \cdot \sum_{g=1}^G 1 \{ \beta_i^{(g)}(t_j) \in BOND_{\beta_i}(t_j) \} \right) \quad (7.6) \end{aligned}$$

, where  $\beta_i^{(g)}(t_j)$  denotes the value of the  $g^{th}$  MCMC sample for fixed effect  $\beta_i$  at grid point  $t_j$ . Subsequently all  $MULTIBONDS_{\beta_i}(t_j) = 0$  are again replaced with  $(2 \cdot G)^{-1}$  and any  $MULTIBONDS_{\beta_i}(t_j) > 1$  are set to 1.

Table 7.1 lists all effect clusters for the two types of SCBs (*type I* and *II*) as well as the posterior probabilities derived from them (MULTIBONDS and SIMBAS) for a subset of three predictors ('c.Linking', 'c.Linking:s.AI' and 's.Word.Length') of the WFMM for electrode AFz. A comparison of the clusters based on probabilities with their respective CI-based counterparts (i.e. *BONDS/SCB I* and *SIMBAS/SCB II*) shows that the two inference bases in fact flag the same time points as significant. Going back to the example of the 'c.Linking' predictor, figure 7.12 illustrates its sole effect cluster for a significance threshold of 0.05 (the complement of the credible band's probability content of 0.95) for the BONDS (figure 7.12a) and SIMBAS (figure 7.12b), which here yield very similar results.

To assess the performance of the different types of posterior probabilities in a somewhat more comprehensive setting, the number of time points classified as significant by all available methods (pointwise CIs, *type I* and *II* SCBs and the respective posterior probabilities)



**Figure 7.12:** a: *Type I* SCBs (blue bands) and BONDS for the 'c.Linking' fixed effect curve (350–950 ms). b: Same but with *type II* SCBs and SIMBAS. In both plots, the broken green and red lines indicate the start/end points of the effect cluster.

were not only calculated for electrode AFz and selected fixed effects, but for the complete ERP model of the 'genitive' ERP-segment, which comprises separate WFMMs for each of 63 electrodes. Including the '(Intercept)' fixed effect, 8 posterior fixed effect curves for each of 63 electrodes with  $T = 350$  were thus classified, amounting to a total of 176400 points. Further, each probability type's 'internal' sensitivity and specificity were computed, taking the respective CI/SCB-based classification as the relevant baseline for comparison: the 'true' underlying positives (significant time points) and negatives (non-significant time points) for the pointwise probabilities were provided by an assessment based on the pointwise CI, and those for the MULTIBONDS and SIMBAS by inference based on the respective SCB. The resulting sensitivity and specificity indices can thus be regarded as a measure of internal consistency between each of the CI/SCB types and its accompanying probability type. Pointwise CIs and SCBs were computed with as 95% CIs/SCBs and the significance threshold was set to 0.05 for all probability types.

Table 7.2 provides the numeric details: pointwise CIs flagged a total of 43421  $t$  as significant (24.6% of all  $t$ ), all of which were also flagged on the basis of pointwise probabilities. Only 5 additional  $t$  were classified as significant based on pointwise probabilities, amounting to a sensitivity of 1 and a specificity of 0.99996.

*Type I* SCBs were considerably more conservative, flagging only 19670  $t$  as significant, 11.2% of all  $t$  and 45.3% of those which were classified as positives by the pointwise credible intervals. According to the corresponding MULTIBONDS, all of the  $t$  identified as significant by the *type I* SCB plus an additional 25 are flagged, yielding an 'internal' sensitivity of 1 and a specificity of 0.99984.



The *type II* SCBs classified 19746 *t* as significant, (11.2% of all points and 45.5% of those flagged by the pointwise CIs). Of these, all were again classified as significant by the SIMBAS, plus 1 not flagged by the *type II* SCBs, resulting in a sensitivity of 1 and a specificity of 0.99999. While the pointwise probabilities and the SIMBAS thus perform slightly better in terms of internal consistency than the MULTIBONDS in the models used for assessment, for most practical purposes they can be considered on a par.

Basis	Positives	Negatives	True +	False +	True -	False -	Sens.	Spec.
pw. CI	43421	132979						
pw. prob	43426	132974	43421	5	132974	0	1	0.99996
SCB I	19670	156730						
BONDS	19695	156705	19670	25	156705	0	1	0.99984
SCB II	19746	156654						
SIMBAS	19747	156653	19746	1	156653	0	1	0.99999

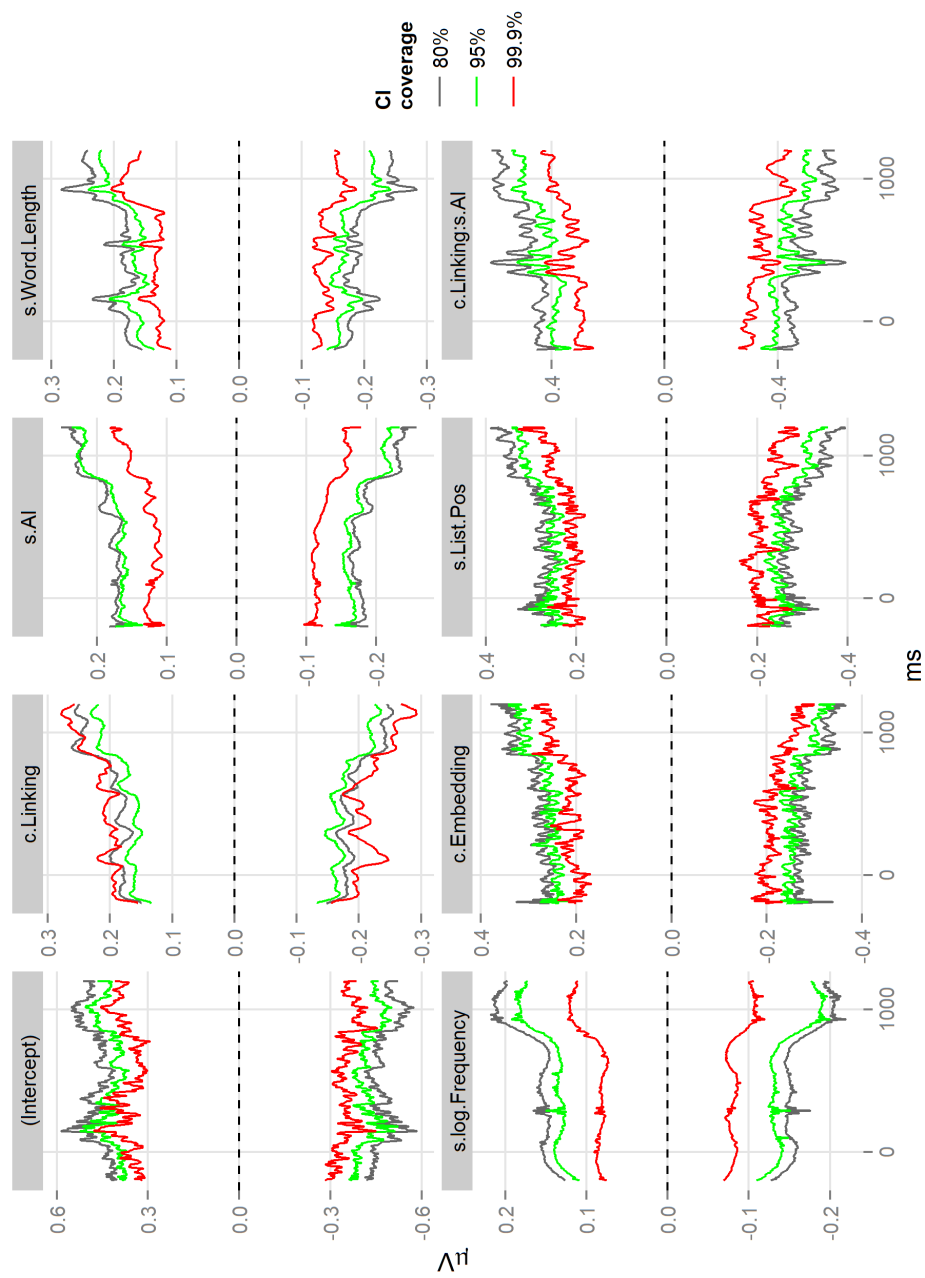
**Table 7.2:** 'Internal' sensitivity and specificity of different posterior probability types in the complete 63-electrode model of the ERP experiment (including the intercept effect). Baseline for pointwise probabilities are classifications as derived by pointwise CIs, for simultaneous probabilities the baselines are provided by classifications based on the type I and II SCBs. 'Positives' are time points flagged as significant, 'Negatives' are time points not flagged as significant.

Having a brief look at the consistency between MULTIBONDS and SIMBAS, a total of 810 *t* were flagged by MULTIBONDS but not SIMBAS and vice versa. This amounts to a disagreement in 0.46% of all *t*.

When interpreting MULTIBONDS, it is key to remember that they directly depend upon the pointwise AND simultaneous credible intervals (as outlined above), since the difference between these is used to construct the functional BOND used as a basis for their computation. This difference between the two types of credible intervals and thus the width of the resulting BOND, in turn, will also be contingent upon their probability level used for the calculation of the CIs.

Figure 7.13 compares the BOND of all fixed effects at electrode AFz at three different probability levels, 80%, 95% and 99.9% and a quick look reveals that one particular sequence is visible with all but the 'c.Linking' predictors: it is illustrated by the 's.Word.Length' and 'c.Linking:s.Al' fixed effect curves, where the *narrowest* BOND is based on the 99.9% and the *widest* BOND on the 80% credible intervals, with the 95% CI-based BOND falling in between. This pattern can further be seen in the remaining effect curves to varying degrees and may reflect a tendency of a decreasing BOND width with increasing target probability of the respective CIs: after all, a pointwise 99.9% CI may already completely contain a very high proportion of all sampled curves, resulting in shorter distances between the respective pointwise and simultaneous credible intervals when compared with CIs at lower probability levels, and thus also a narrower functional BOND.

Table 7.3 exemplifies the consequences of this pattern for MULTIBONDS-based inference, listing all effect clusters for electrode AFz based upon MULTIBONDS for the three different credible band-probability levels with varying significance thresholds  $\alpha$  (80% SCBs:  $\alpha = 0.2$  and 0.05; 95% SCBs:  $\alpha = 0.05$  and 0.001; 99.9% SCBs:  $\alpha = 0.001$ ). A comparison of the clusters based upon the same significance threshold, but with MULTIBONDS computed



**Figure 7.13:** Comparison of BONDs of all fixed effects at electrode AFz for different CI/SCB levels (80%, 95% and 99.9%).

Fixed Effect	Sign. thresh.	CI prob.	Cluster	Start	End	Length
c.Linkig	0.2	80%	1	400	884	484
c.Linkig			2	1108	1200	92
c.Linkig:s.AI			1	504	780	276
c.Linkig:s.AI			2	1196	1200	4
c.Linkig	0.05	80%	1	440	856	416
c.Linkig:s.AI			1	524	644	120
c.Linkig:s.AI			2	696	732	36
c.Linkig			1	432	860	428
c.Linkig:s.AI	0.05	95%	1	520	652	132
c.Linkig:s.AI			2	688	748	60
c.Linkig			1	548	824	276
c.Linkig			1	552	820	268
c.Linkig:s.AI	0.001	95%	1	548	580	32
c.Linkig	0.001	99.9%	1	548	580	32
c.Linkig:s.AI	0.001	99.9%	1	548	580	32

**Table 7.3:** Details for effect clusters based on BONDS-inference for three different CI-coverage levels (80%, 95% and 99.9%) and varying significance thresholds.

from different CI-levels (i.e.  $\alpha = 0.05$  with 80% and 95% credible bands;  $\alpha = 0.001$  with 95% and 99.9% credible bands) shows that some estimates are in fact *more conservative* for the clusters based upon the *less conservative* SCBs, with single clusters vanishing, shortened or broken up into several shorter clusters. This is generally true for the 'c.Linkig:s.AI' predictor, since the respective BONDS follow the pattern described in the previous paragraph (see figure 7.13); it also applies for the comparison of the 'c.Linkig' cluster at the  $\alpha = 0.05$  probability threshold computed from 80% and 95% SCBs, since the BONDS for the former are broader than those for the latter. The reverse pattern of *less conservative* estimates for the clusters based upon the *less conservative* SCBs can be seen when comparing the 'c.Linkig'-related cluster at the  $\alpha = 0.001$  probability threshold computed from 95% and 99.9% SCBs: since the BONDS related to the former are actually narrower than those associated with the latter, the resulting cluster is slightly longer when the MULTIBONDS are computed from the 95% SCBs. Note that such behaviour also implies that the internal consistency between SCB-based and MULTIBONDS-based inference may deteriorate when choosing non-default significance thresholds for MULTIBONDS inference, i.e. when the threshold is not set at the complement of the respective SCB's probability level.

While the outline and assessment presented here is rather anecdotal and informal in nature, I hope that it will be helpful in using MULTIBONDS for inference, considering their experimental character. Ultimately, the shape and width of any BOND and thus the values of the corresponding MULTIBONDS will depend on the form of the underlying posterior and may reflect different local and global properties of a functional posterior estimate. When using any of the multiplicity-adjusted probability types presented here for inference, it is thus important to be aware of their interpretation and their experimental character until a more comprehensive discussion of their properties under different conditions is supplied.

## 7.8 Summary

This chapter has introduced the *wrapfmm* package for the R statistical computing environment (R Development Core Team, 2013), which provides a wrapper for the WFMM software of Morris and Carroll (2006). By making use of *lmer*'s mixed model formula interface for WFMM specification with the *wfmm* function, it offers a high-level mixed model interface and makes transition towards using the WFMM easy for users already familiar with the *lme4* package of Bates et al. (2014). *wrapfmm* further focuses on tools enabling users to explore, mine and plot the functional data (*plot*) as well as a WFMM's posteriors (*plot*, *summary*, *follow\_up*) with a high degree of flexibility and abstraction, thus maximising the potential for creative handling of data and posteriors while minimising the amount of code required.

For inference about fixed effects, the *wrapfmm* interface offers a number of different bases, including pointwise CIs and SCBs as well as corresponding posterior probabilities. Section 7.7.2 presented two instances of a novel concept I here referred to as 'simultaneous probabilities' in more detail: the SIMBAS developed by Meyer et al. (2015) and the MULTIBONDS, which are an original contribution of this thesis. Both allow to base inference about functional posterior estimates on continuous posterior probabilities which take multiplicity along the grid  $t$  into account and are derived from Bayesian SCBs. While the two probability types are based on different algorithms, both provide a means for inference using multiplicity-adjusted probabilities without defining (often arbitrary) minimum effect sizes or using brute-force corrections – in fact, the adjustment falls out directly from the SCBs, which are in turn a natural 'product' of the underlying functional posteriors.

I also pointed out a number of limitations, some of which stem from the core WFMM software, others from the package itself, including the experimental character of some of the package's features, such as the MULTIBONDS introduced above. Implications of these and other limitations as well as possible extensions will be discussed in more detail in chapter 9.

## Chapter 8

# The *stepmom* package

### 8.1 Aims and scope

The *stepmom* (**S**patiotemporal **E**lectrophysiological **M**odel **M**aps) package builds upon and extends *wrapfmm* for mixed-model analyses of multi-channel ERP data. Using the WFMM of WFMM of Morris and Carroll (2006) as the underlying mixed-model engine, the analysis can be performed without specification of temporal windows and spatial regions of interest.

Generally, *stepmom* pursues the aim of providing the same degree of flexibility and usability as *wrapfmm* for the more complex case of full-fledged multi-electrode ERP data sets. In addition, the user-interfaces of *stepmom* functions copy or closely follow those of the *wrapfmm* functions – familiarity with *wrapfmm* will thus considerably ease starting to work with *stepmom*. In fact, *stepmom* offers nearly all the functions of *wrapfmm* with appropriate methods for modelling, plotting and exploring multi-channel ERP data on a by-electrode level. Going beyond the electrode-level, *stepmom* provides a number of additional functions for plotting ERP data as well as model output as scalp maps. To do justice to the full spatiotemporal nature of ERPs, data and WFMM-based modelling results can further be displayed as interactive and fully controllable scalp map animations over time.

Key features of *stepmom* thus include:

- Methods for performing *wfmm*- (or *lmer*-) based analysis of ERP data via the familiar *lme4* mixed-model interface.
- Full availability of key features of *wrapfmm* (e.g. the different inference bases presented in section 7.7).
- Appropriate methods of functions from *wrapfmm* for exploring ERP data and WFMM results on the electrode-level.
- Functions for plotting ERP data (*vom* methods – ‘voltage maps’) as well as model output (*mom* methods – ‘model maps’) as static plots or as HTML-embeddable, interactive

animations (*ani*), all offering a high degree of flexibility and control over graphics parameters for users.

## 8.2 Availability & compatibility

As pointed out in the preface, the asynchrony of the package versions with current versions of R limits their practical value as presented here. However, they can be read as (extended) supplementary material to the ERP-analyses presented in section 4.2.6 of the first part and as complementary background chapters on new packages that will be available soon, retaining almost all of the basic functionality presented here, while at the same time offering many more features; for more information, including the status of the packages as presented here, please refer to the preface and the introductory comments to 6.

## 8.3 Implementation

As *wrapfmm*, the *stepmom* package is written in pure R code and its objects, functions and classes are implemented within R's *S3* system (e.g. Matloff, 2011, chapter 9). All comments provided in section 7.3 of the *wrapfmm* introductory chapter equally apply to *stepmom*.

## 8.4 WFMM backend

The *stepmom* package builds upon the *wrapfmm* package presented in the previous chapter and also requires the WFMM executable installed as computational backend (see section 7.4 for further information).

## 8.5 Limitations

On the modelling level, the most obvious limitation of *stepmom* is its ignorance of the spatial axis of the object structure underlying ERP data. The *wfmm* method implemented for ERP modelling makes use of the resources available in the *wrapfmm* package, i.e. WFMM-based modelling of 1d curve data. Thus, it is currently limited to modelling each electrode's data independently, thus ignoring the spatial correlations between the multiple electrodes. This issue is further discussed in section 9.2.1 below. Note, however, that the new packages currently developed will offer the possibility to fit spatially correlated functional mixed models, addressing this shortcoming (see also the comments in the preface). The version presented here only features built-in support for electrode setups making use of electrodes included in

the 10-10 system. As far as interaction with the WFMM core software are concerned, all remarks on limitations of the *wrapfmm* package apply equally.

## 8.6 Overview

### 8.6.1 Getting oriented

The following is a brief list of the most important functions the package comprises. They are arranged in the 'order of appearance', reflecting the sequence of application within a typical workflow.

**mom\_base** sets up a `momBase` object which holds the matrix with the ERP segment data, a data frame with variables reflecting related information for each trial included as well as associated information about the data and the montage used.

**plot** allows to plot ERP data for selected electrodes as curves, providing an easy-to-use interface for breaking up the data by the levels of multiple factors, including on-the-fly conversion of numerical variables to factors and computation of difference curves. It is a generic function with methods for objects of class `momBase`, `wfmMom` and `postWfmMom`. When applied to a `wfmMom` or `postWfmMom` object, the data plots can be combined with visualisations of posterior estimates of fixed effect curves. Its versatility makes this function a useful tool for exploring curve data prior to analysis and visualising a WFMM's output together with the underlying data.

**vom** (for 'voltage map') plots multi-channel ERP data as voltage scalp maps. As *stepmom*'s `plot` function, `vom` has a powerful high-level user interface providing numerous options for flexible specification of voltage maps, including on-the-fly computation of averages for arbitrary time windows as well as calculation of difference maps, plotting the position of electrodes and full control over numerous graphics parameters.

**wfmm** is *stepmom*'s core function and is used for specifying and running by-electrode WFMMs for multi-channel ERP data without specification of temporal windows and spatial ROIs. It is applied to a `momBase` object, the corresponding model structure is easily defined via an *lme4* mixed model statement (Bates et al., 2014) and options for the core WFMM executable can be specified conveniently. As the corresponding method of the *wrapfmm* package, it takes care of the complete R/WFMM input/output pipeline, generates the same default inferences (see section 7.7) and returns an object of class `wfmMom`.

**lmerMom** models time-window mean ERP data using the `lmer` function of package *lme4* (Bates et al., 2014). After specification of time windows of interest, the function automatically computes mean values for each window and electrode and fits separate LMMs to each time window/electrode data set. Outliers can be trimmed before

model fitting and `lmerMom` further offers a nearest-neighbour  $p$ -value correction based on spatial proximity of electrodes. It returns an object of class `lmeMom` and modelling results can be plotted as scalp maps using the `mom` function.

`summary` returns a curve-plot summary of a `wfmMom` object's posterior fixed or random effects curves or its associated variance components for selected electrodes. For fixed effects, `summary` may also return a data frame with information for each significant cluster. Inference about fixed effects can be based on the same range of inference bases as available with the `wrapfmm` package. The `summary` function is a generic function with methods for `wfmMom` and `postWfmMom` objects, and thus also allows to summarise post-hoc posteriors computed with the `follow_up` function. It can further be used to return a summary plot of a `wfmMom` object's variance components for electrodes of interest.

`mom` (for 'model map') is the montage-level analogue of the `summary` method and plots the results of `wfmm-` (with appropriate methods for `wfmMom` and `postWfmMom` objects) or `lmer`-based analyses of multi-channel ERP data as scalp maps. Numerous class-specific output statistics can be selected for plotting, such as (posterior) mean estimates and probabilities,  $F$ -values, etc., for fixed effects summary maps and variances and SDs for variance component maps. For `wfmm`-based model maps, model results can be averaged within arbitrary time windows. `mom` again offers a user-friendly interface and full control over numerous graphics parameters.

`ani` produces animations of data-based voltage maps or `wfmm`-based ERP model maps using the `animation` package (Xie, 2013), offering the same range of features as `vom` and `mom` (again with methods for `wfmMom` and `postWfmMom` objects). The resulting animations can be embedded in HTML pages and can be controlled interactively via control panels.

`follow_up` is used to compute post-hoc posteriors from a `wfmMom` object's fixed effect posteriors, such as simple effects. The function returns an object of class `postWfmMom` and all of the functions applicable to `wfmMom` objects have appropriate methods for `postWfmMom` objects. The `follow_up` function can also be used to compute 'higher order' post-hoc posteriors from post-hoc posteriors contained in another `postWfmMom` object.

`chain_plot` is a generic function for making (informal) diagnostic plots for MCMC chains of fixed effect posteriors for an electrode of a `wfmMom` or `postWfmMom` object. Plot types comprise chain plots, density plots and boxplots.



## 8.6.2 A quick tour

The current section will give a brief tour of *stepmom*'s core functions, providing code examples and pointing out relevant features. Functions which have corresponding methods in the *wrapfmm* package (e.g. `plot` and `summary`) will mostly be illustrated very cursorily.

### 8.6.2.1 The data set

The data used for illustration purposes comes from the same data set used for demonstration of the *wrapfmm* functions in the previous chapter. While the latter represents a subset of eight subjects and a single electrode (Pz), the data set used in the following features the data of 25 electrodes. For each subject, single-trial ERPs were averaged within each of the four cells obtained by the combination of all levels of the two binary factors 'Predictability' and 'Session'. More information about the data is provided in section 7.6.2.1 in the preceding chapter.

### 8.6.2.2 `mom_base`: setting up a base object

The *stepmom* equivalent of `wfmm_base` (section 7.6.2.2) is the `mom_base` function. In addition to the three-dimensional array holding the multi-channel ERP data and information about the corresponding variables and associated units, the `momBase` object returned contains a data frame providing information about the underlying electrode montage. As pointed out in the section on current limitations (8.5 above), the version presented here only offers built-in support for electrode montages comprising electrodes of the 10-10 system.t.

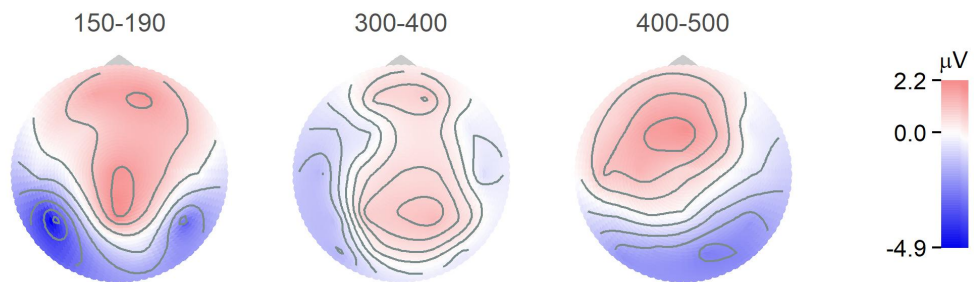
The following call constructs the base object:

```
n400.avg.base <-  
  mom_base(Y = n400.avg.data,  
           variables = n400.avg.vars,  
           units = list(x = "ms",  
                        y = expression(paste(mu, "V"))))
```

### 8.6.2.3 `vom`: plotting ERP data as voltage scalp maps

With the `mom_base` object set up, the ERP data can be explored as voltage scalp maps using the `vom` function. Time windows for averaging and plotting can be defined arbitrarily, including non-adjacent intervals and/or intervals of different lengths. The below call produces figure 8.1 and shows the overall mean voltage maps for an early 40 ms time window in the range of the N170 component (150–190 ms) and two later 100 ms windows in the time range of the N400 component.

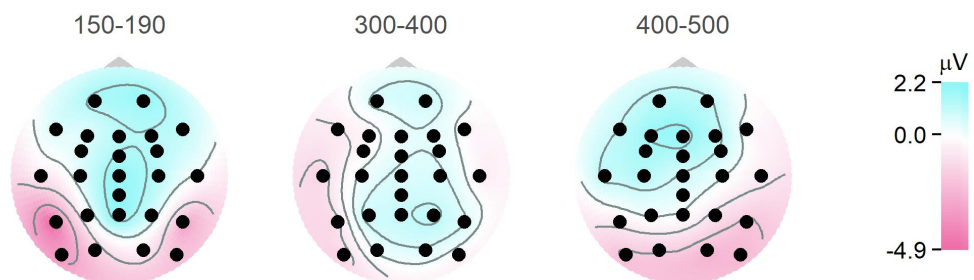
```
vom(n400.avg.base,
    time.windows = list(c(150, 190),
                        seq(300, 500, by = 100)))
```



**Figure 8.1:** Example voltage maps for three different time windows produced with the `vom` function.

`vom` allows users to customise the voltage maps in numerous ways; the next call defines funky map colours, plots all electrode positions of the montage used and reduces the number of contour lines to five (see figure 8.2).

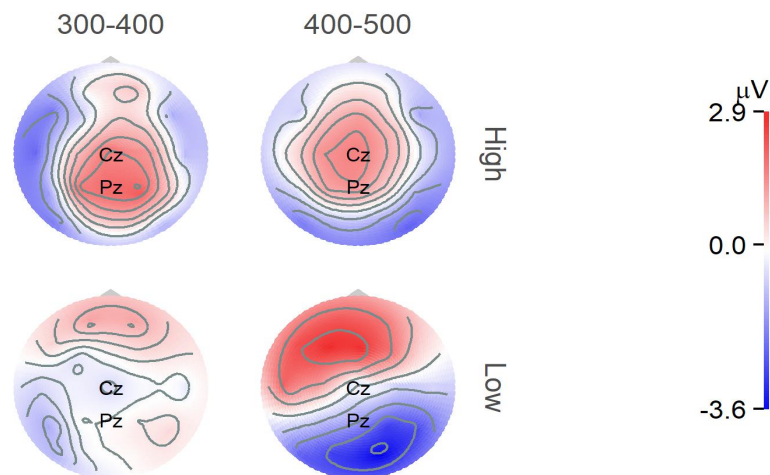
```
vom(n400.avg.base,
    map.colours = c("hotpink2", "white", "cyan2"),
    montage = list(show = "all"),
    contours = list(bins = 5),
    time.windows = list(c(150, 190),
                        seq(300, 500, by = 100)))
```



**Figure 8.2:** Same as in previous figure, but with alternative map colours and all electrodes shown.

As *wrapfmm*'s and *stepmom*'s summary and plot methods, *vom* makes use of the powerful facetting interface of the *ggplot2* package (Wickham, 2009). Using `facet.grid` or `facet.wrap`, the data can be broken up by any combination of factor levels; continuous numeric variables can be converted to factors on the fly using the `factor.options` argument. Below, separate voltage maps for the two levels of the 'Predictability' factor are defined for each of the two later time windows within the N400 time range, plotting the names of electrodes Cz and Pz (figure 8.3):

```
vom(n400.avg.base,
    facet.grid = Predictability ~ time.windows,
    montage = list(show = c("Cz", "Pz"),
                   shape = "name"),
    time.windows = list(seq(300, 500, by = 100)))
```



**Figure 8.3:** Mean voltage maps for each level of the 'Predictability' factor for two 100 ms time windows.

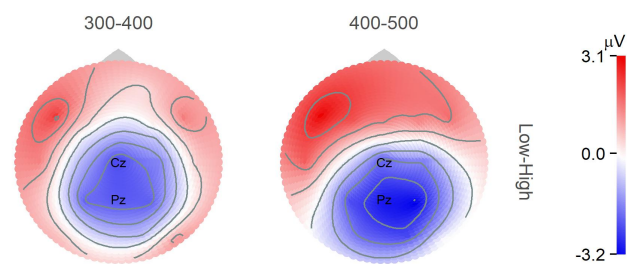
Making use of the `diffmaps` argument, difference maps between specified factor levels can be computed and plotted. Specification of difference maps works parallel to the definition difference curves with *wrapfmm*'s and *stepmom*'s summary function via its `diffcurves` option: the following call yields figure 8.4, which plots the difference between the 'Low' and 'High' levels of 'Predictability' shown separately in figure 8.3.

```
vom(n400.avg.base,
    diffmaps = list("Low-High" = list(factor = "Predictability",
                                      difference = c("Low", "High"))),
```

```

        plot.diffmaps.only = TRUE),
facet.grid = Predictability ~ time.windows,
montage = list(show = c("Cz", "Pz"),
                shape = "name"),
time.windows = seq(300, 500, by = 100))

```



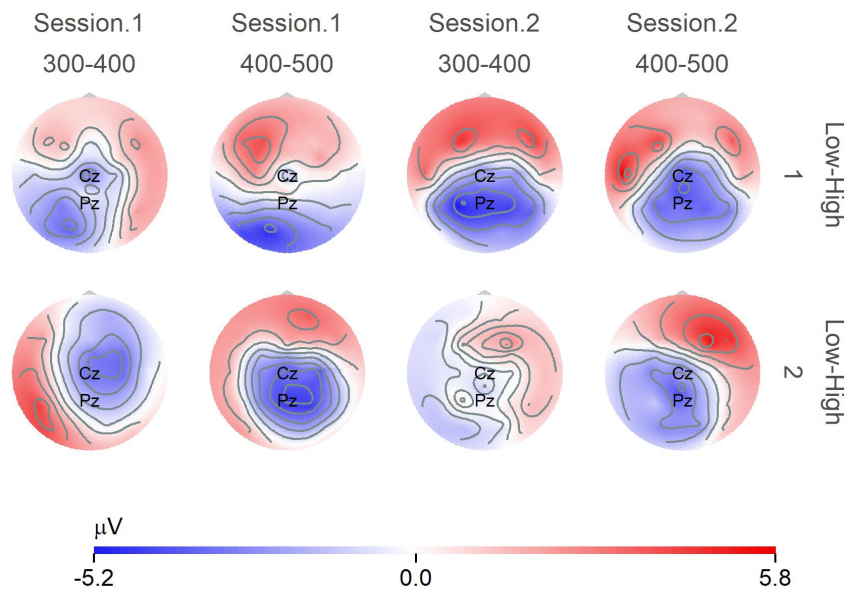
**Figure 8.4:** 'Low' minus 'High' predictability difference maps for two 100 ms time windows.

Finally, the following code snippet combines facetting and subsetting to produce 'Predictability' difference maps of two individual subjects separately for each session (figure 8.5), illustrating intraindividual variability across sessions as well as interindividual variability.

```

vom(n400.avg.base,
  facet.grid = Subject + Predictability ~ Session + time.windows,
  subset = .(Subject %in% 1:2),
  diffmaps = list("Low-High" = list(factor = "Predictability",
                                     difference = c("Low", "High")),
                  plot.diffmaps.only = TRUE),
  montage = list(show = c("Cz", "Pz"),
                  shape = "name"),
  time.windows = seq(300, 500, by = 100))

```



**Figure 8.5:** Difference maps for two individual subjects, split by 'Session'.

Note that `vom` internally performs the interpolation of values to the complete grid by applying the `gam` function of the `mgcv` package (Wood, 2006) with a spherical spline. The default settings may not be optimal for all applications and can be adapted when calling `vom` via the `gam.settings` argument. Also, in some cases the parameters for the interpolation grid (such as its resolution) and other parameters may have to be adjusted to reduce the size of the resulting plot or to optimise other graphics parameters.

#### 8.6.2.4 `wfmm`: fitting wavelet-based functional mixed models to multi-channel ERP data

The interface of *stepmom*'s `wfmm` method largely parallels that of *wrapfmm*'s `wfmm` function. The only additional relevant argument which it may take is `electrodes`, which serves to specify a subset of electrodes for modelling. Please refer to section 7.6.2.3 about *wrapfmm*'s `wfmm` method for more detailed information – all comments found there equally apply to *stepmom*'s `wfmm` function!

Once called, `wfmm` will launch one separate WFMM for each of the electrodes included. Currently, it tries to find a compromise between running multiple of these separate WFMMs in parallel while not overtaxing computational resources: a first WFMM is launched and after the initial, potentially rather memory-intensive preprocessing phase, a second one is launched. Once one of these is finished, the next one is launched, and so on. This way, the separate WFMMs are run in a roughly cascaded fashion. While running, `wfmm` will create one

subfolder for each electrode within the main model directory where all of the WFMM output files are stored.

The call below specifies the same model as defined in the previous chapter for the single-electrode WFMM: it includes an intercept fixed effect as well as effects for the 'c.Session' and 'c.Predictability' variables plus their interaction in the fixed effects part of the model. In addition, it again features by-subject random intercepts and by-subject random slopes for the 'c.Predictability' effect in the random effects part.

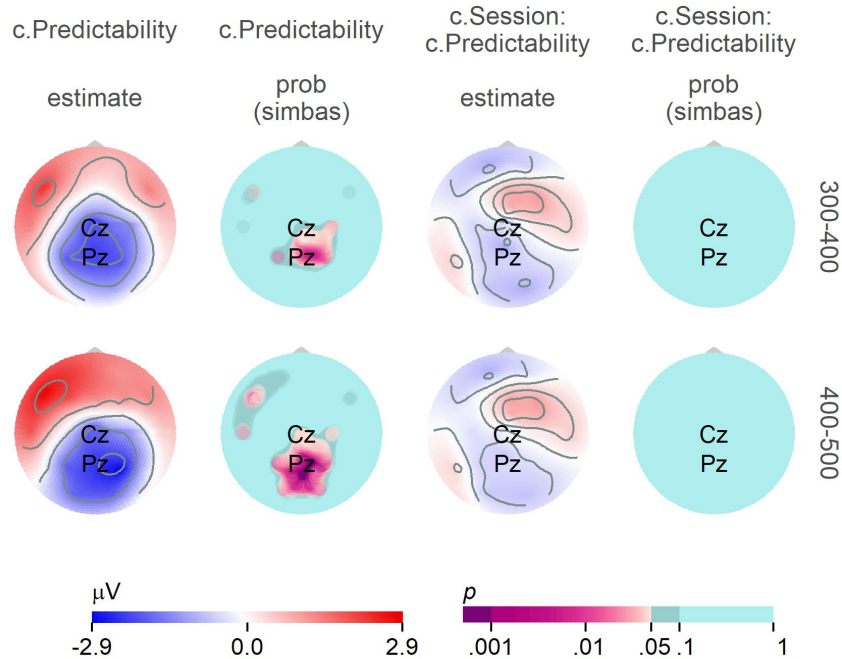
```
n400.avg.wfmm <-  
  wfmm(x = n400.avg.base,  
        disk.id = "n400.avg.wfmm",  
        formula = ~ 1 + c.Session * c.Predictability +  
          (1 + c.Predictability | Subject),  
        PostProcessSpecs = list(compute_U = 1),  
        MCMCSpecs = list(B = 2000, burnin = 1000, thin = 3),  
        wavespecs = list(wavelet = "sym8"))
```

#### 8.6.2.5 mom: summarising a WFMM's posterior estimates as 'model maps'

The `mom` function is the scalp-level analogue of the `summary` method: it plots the output of a `wfmm`-based (or `lmer`-based) ERP analysis as scalp maps using a user interface which incorporates elements of the `vom` and `summary` functions. It requires specification of (arbitrary) time windows within which the output statistics selected for plotting (e.g. the posterior mean estimate and/or posterior probabilities) are averaged. In the following, summary model maps of the WFMM defined in the previous section are plotted, concentrating on the estimates of the 'c.Predictability' and 'c.Session:c.Predictability' fixed effects for two 100 ms time windows. If not specified otherwise, `mom` selects the posterior mean estimate and the SIMBAS (see section 7.7.2 for details) as statistics for display.

```
mom(n400.avg.wfmm,  
    ffx.settings = list(predictors = c("c.Predictability",  
                                       "c.Session:c.Predictability")),  
    facet.grid = time.windows ~ predictor + statistic,  
    time.windows = seq(300, 500, by = 100),  
    montage = list(show = c("Cz", "Pz"),  
                   shape = "name"),  
    legend = list(position = "bottom"))
```

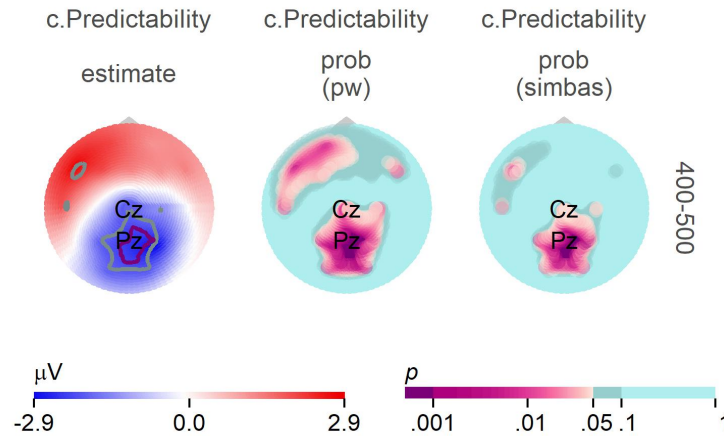
The resulting maps in figure 8.6 show significant clusters for the 'c.Predictability' effect in both time windows, one posterior cluster with a negative slope and a maximum around electrode Pz as well as smaller left anterior clusters with inverted polarity (positive slopes). The posterior clusters correspond to a classical N400 effect. For the interaction effect, the maps do not indicate any significant clusters.



**Figure 8.6:** Model scalp maps for the posterior mean estimates and associated SIMBAS for the 'c.Predictability' and 'c.Session:c.Predictability' fixed effects.

As inference bases for wfm-based models, *mom* offers the posterior probability types also available with the *summary* method, as detailed in section 7.7.2 and multiple probability types may be selected for display in the same plot. In figure 8.7 below, the posterior mean estimate and the corresponding SIMBAS are again selected and the less conservative pointwise probabilities are added. Further, the contours plotted within the maps of the posterior mean estimates are specified to demarcate those regions which are below the respective significance threshold as defined by the SIMBAS – this allows exact identification of the extent of the respective effects within the estimate maps and immediately shows the direction of the effect. Note that in contrast to the previous plot, figure 8.7 only shows the maps for the later 400–500 ms time window

```
mom(n400.avg.wfmm,
    ffx.settings = list(predictors = c("c.Predictability"),
                        statistics = c("estimate", "simbas", "pw.prob")),
    contours = list(project.p = "simbas"),
    facet.grid = time.windows ~ predictor + statistic,
    time.windows = c(400, 500),
    montage = list(show = c("Cz", "Pz"),
                   shape = "name"),
    legend = list(position = "bottom"))
```



**Figure 8.7:** Model maps for the posterior mean estimates, associated SIMBAS and pointwise probabilities for the 'c.Predictability' fixed effect. Contours within the estimate maps demarcate regions flagged as significant by the SIMBAS.

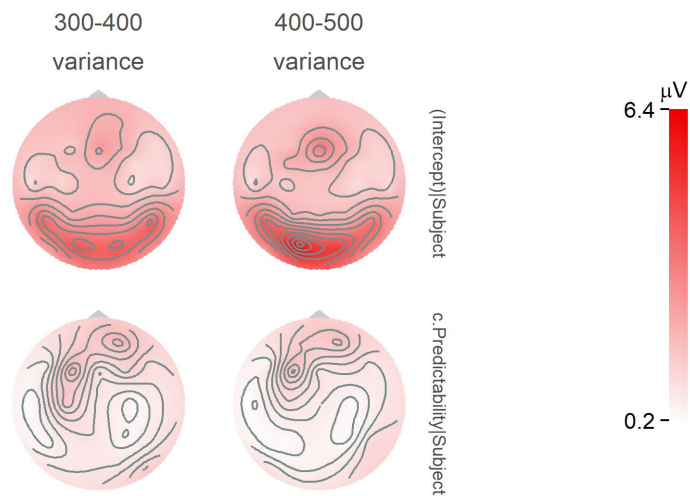
As the plot shows, the SIMBAS are clearly more conservative than the pointwise probabilities. The joint display of both multiple probability types provides an additional source of information about the reliability and extent of the effects when accounting for multiplicity along the temporal axis (or not).

Finally, figure 8.8 is the result of the next call to `mom`, which summarises the variances of the by-subject random intercept and random slope for the 'c.Predictability' effect for the same two 100 ms time windows used above for the fixed effects model maps in figure 8.6. For the random intercept, the largest variances are seen in occipito-temporal electrodes, while the variances are highest at frontal electrodes and lowest at temporal and parietal sensors for the random slope.

```
mom(n400.avg.wfmm,
    type = "VC",
    facet.grid = VC ~ time.windows + statistic,
    vc.settings = list(VC = c("(Intercept)|Subject",
                             "c.Predictability|Subject")),
    time.windows = seq(300, 500, by = 100))
```

The `mom` function allows users to produce useful summary maps of ERP modelling results. However, the resulting maps reflect averages of estimates and probabilities within the time windows defined – depending on the length of the time windows used for the model maps and their degree of overlap with the actual temporal extent of the effects, they may or may not represent the effects' characteristics appropriately. Ill-chosen temporal windows may





**Figure 8.8:** Variance maps for the by-subject random intercepts and random slopes for the 'c.Predictability' effect.

not capture the central characteristics of an effect or miss it altogether. Thus, whenever possible, model maps should be presented together with electrode-level summary plots showing the exact temporal extent and dynamics at selected electrodes as well as with animated scalp maps which convey the exact spatial and temporal extent of effects at finer temporal resolutions, down to the original temporal sampling rate used for the analysis. The two functions used for these complementary summaries are presented in the next two sections.

#### 8.6.2.6 `ani`: animated voltage or model maps

The `ani` generic function produces animations of data-based voltage maps or `wfmm`-based model maps with Xie's (2013) *animation* package. Its output is a basic HTML page containing the animation as well as directories containing the single animation frames and additional files. The pages feature control panels which allow interactive control of the animations, including buttons for forwarding/rewinding, pausing and adjusting presentation speed. The code can be easily inserted into other HTML pages, such that multiple animations can be commented and presented together. This way, the data underlying an ERP experiment as well as the results of the `wfmm`-based modelling can be made available as supplementary material and readers can delve into either on their own and may explore time ranges, spatial regions or effects which were not the focus of the respective study in more detail. Thus, `ani` offers a way to convey a maximum amount of information about an ERP data set efficiently and interactively. `ani` makes use of the same user interfaces as `vom` and `mom`; it only needs additional information

about page and file names and allows to set a number of options. As with `vom` and `mom`, time windows can be specified freely and the temporal resolution may be lowered down to the original sampling rate. The following call will produce animations for the estimates for the 'c.Predictability' and 'c.Session:c.Predictability' fixed effects spanning the same time range as used in the model maps in figure 8.6 at a resolution of 8 ms, writing the output into the current R working directory.<sup>1</sup>

```
ani(n400.avg.wfmm,
    ffx.settings = list(predictors = c("c.Predictability",
                                      "c.Session:c.Predictability")),
    facet.grid = statistic ~ time.windows + predictor,
    time.windows = seq(300, 500, by = 8),
    save.settings = list(img.name = "ani_mom_demo"),
    ani.settings = list(title = "ani mom demo",
                        imgdir = "ani mom demo",
                        outdir = getwd(),
                        htmlfile = "ani_mom_demo.html",
                        autoplay = FALSE))
```

#### 8.6.2.7 summary and plot: summarising a WFMM's posterior estimates and plotting the underlying data for selected electrodes

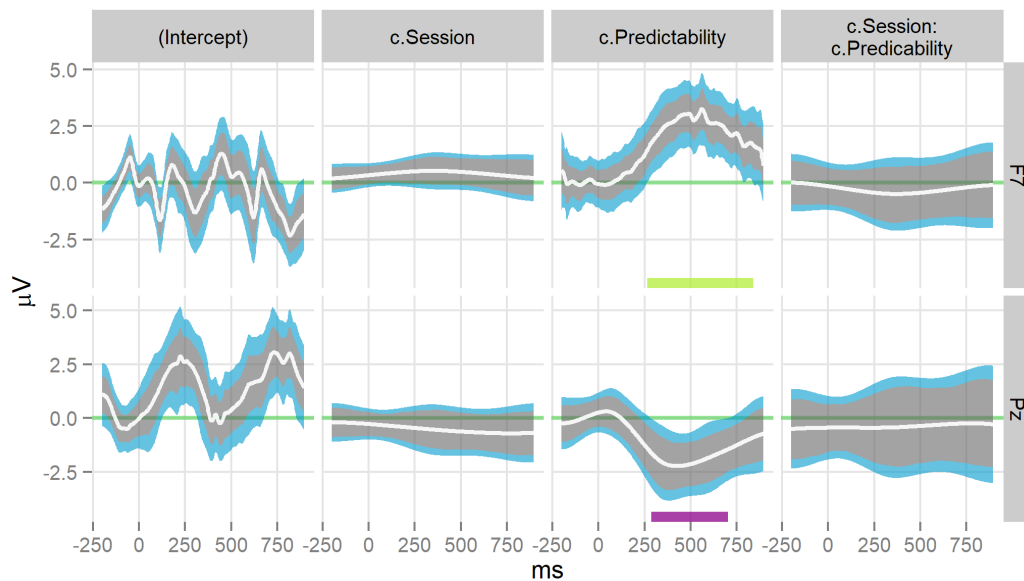
*stepmom*'s summary and plot methods work parallel to the *wrapfmm* equivalents, offering one additional electrodes argument which allows selection of one or more electrodes for plotting. Here, I will only present one example call and plot for each method – more examples as well as more detailed information can be found in sections 7.6.2.4 and 7.6.2.6 in the *wrapfmm* introductory chapter. The help page of the `summary.wfmmMom` method lists further example calls, including summaries of random effects and variance components. For more plot examples, see the help page of `plot.wfmmMom`.

The call below returns figure 8.9, which summarises the posterior estimates for the model's fixed effects at two electrodes (F7 and Pz) with inference based on *type II* SCBs. Note that the 'electrode' factor has to be included as a term in the facetting formula and that the order of the electrode names within the `electrodes` vector determines the order in which the panels are eventually arranged in the plot.

```
summary(n400.avg.wfmm,
        facet.grid = electrode ~ predictor,
        title = NULL,
        electrodes = c("F7", "Pz"))
```

---

<sup>1</sup>Examples of animated model maps are available as supplementary materials at <http://doi.org/10.18452/19291>; see also section 4.2.6 for more information.



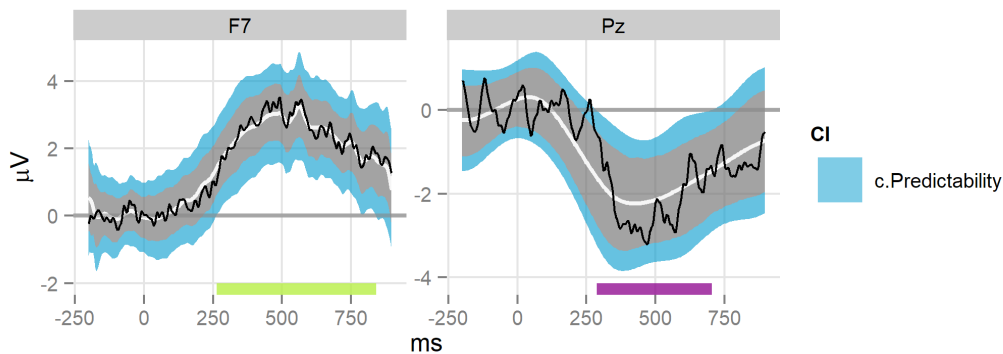
**Figure 8.9:** Posterior fixed effects summary for two electrodes (F7 & Pz) with default settings. Assessment of significance is based on whether the *type II* SCBs (outer blue bands) contain zero or not.

In the following, the 'Low' minus 'High' predictability difference curves are plotted along with the posterior estimates of the 'c.Predictability' fixed effect for the same two electrodes (figure 8.10). The 'electrode' factor is again used a facetting term to make sure each electrode gets its own panel.

```
plot(n400.avg.wfmm,
     diffcurves =
       list("Low-High" = list(factor = "Predictability",
                              difference = c("Low", "High")),
            plot.diffcurves.only = TRUE),
     ci = list('c.Predictability' = list()),
     ci.type = c("pw", "sim.II"),
     facet.wrap = ~ electrode,
     electrodes = c("F7", "Pz"))
```

### 8.6.2.8 follow\_up: computing post-hoc posteriors

The interface of *stepmom*'s `follow_up` method is again fully parallel to that of its *wrapfmm* analogue (see section 7.6.2.5 for more details). When applied to a *wfmMom* or *postWfmMom* object, it calculates the specified post-hoc posteriors for all or a subset of electrodes. The



**Figure 8.10:** 'Low' minus 'High' predictability difference curves with posterior estimates of the 'c.Predictability' fixed effect for two electrodes (F7 & Pz). Assessment of significance is based on *type II* SCBs (outer blue bands).

following example computes the posterior cell means estimates for 'High' and 'Low' predictability words:

```
n400.avg.postwfmm <-
  follow_up(n400.avg.wfmm,
    disk.id = "n400.avg.postwfmm",
    formulas =
      list(High = .(beta((Intercept)) +
        beta(c.Predictability) * min(pred(c.Predictability))),
        Low = .(beta((Intercept)) +
        beta(c.Predictability) * max(pred(c.Predictability)))))
```

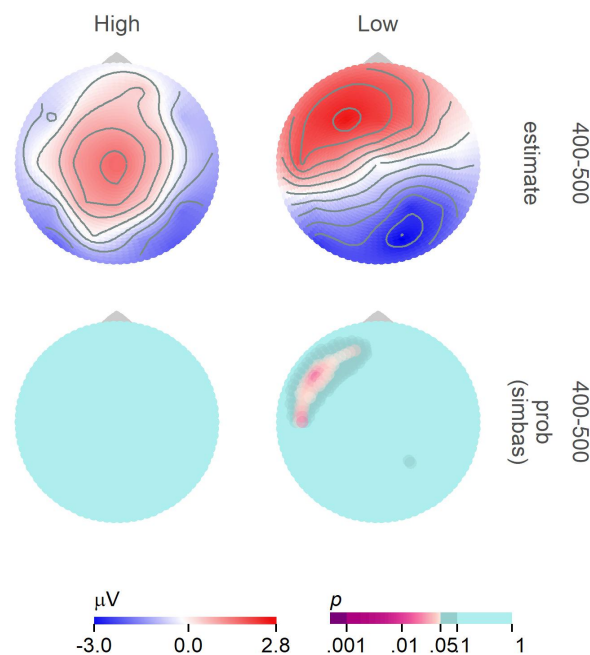
`follow_up` returns an object of class `postWfmMom` with the same range of methods as for `wfmMom` objects, including `summary`, `plot`, `mom` and `ani`. Figure 8.11 shows the model maps for the computed follow-up posteriors for one 100 ms time window, which are the result of the subsequent `mom` call.

```
mom(n400.avg.postwfmm,
  facet.grid = predictor ~ time.windows + statistic,
  time.windows = c(400, 500),
  legend = list(position = "bottom"))
```

As with `wrapfmm`'s `follow_up` method, 'higher-level' post-hoc posteriors can be computed in two ways: either by applying `follow_up` to a `postWfmMom` object, or by accessing previously defined follow-up posteriors 'on the fly' within one and the same function call.

#### 8.6.2.9 `lmerMom`: fitting LMMs to time window average data

The `lmerMom` function fits scalar LMMs to ERP data averaged within time windows using the `lmer` function of the `lme4` package of Bates et al. (2014) and is intended to complement



**Figure 8.11:** Model maps of the posterior cell means estimates for low and high predictability words computed with the `follow_up` function (inference based upon SIMBAS).

`wfmm`-based ERP analyses. As pointed out above (see sections 8.5 and 7.5), application of the WFMM is currently subject to a number of inherent and practical limitations: when compared to a standard LMM fitted with `lmer`, for example, a WFMM currently does not allow to include correlation parameters between random effects. Also, in practice the complexity of a WFMM may be restricted due to limited computational resources, precluding exhaustive modelling of the design structure underlying the data – e.g. for experimental designs requiring multiple by-subject random effects as well as one or more by-item random effects, inclusion of all random factors may not be feasible if any of the random grouping factors has a large number of levels.

In such situations, additional modelling of the ERP data with more complex model structures using `lmerMom` represents a pragmatic complementary feature extraction approach to `wfmm`-based analyses. Following modelling using the WFMM-approach, significant clusters identified by the WFMM can be used for defining a number of time windows for `lmer`-based analysis. The choice of start and end time points for these time windows may be based on start/end times of clusters of interest or around the times at which effects peak (e.g. using electrodes at which relevant effects are maximal). Convergence of `wfmm`- and `lmerMom`-based models for the selected time windows can subsequently be assessed by comparing the respective results. The relative impact of adding/removing random effects, for example, may be checked by plotting and informally comparing model maps of `lmer`-based models of different

complexity.

One prominent issue with `lmer` is the lack of  $p$ -values for fixed effects in its output, which is due to issues with determining the degrees of freedom for computation (see Baayen et al., 2008, for discussion). `lmerMom` computes  $p$ -values based upon upper- and lower-bound degrees of freedom using the `pamer.fnc` function of the *LMERConvenienceFunctions* package (Tremblay & Ransijn, 2013). As an experimental feature, it further offers the possibility of performing multiplicity corrections of  $p$ -values in a way which adapts to the local density of electrode montages: applying the function `p.adjustSP` of package *spdep* (Bivand, 2013) it allows to correct the  $p$ -values of a given electrode within a specific time window against a subset of *local nearest neighbour* electrodes, rather than against the complete set of electrodes of the montage. Such a nearest-neighbour correction can be performed using a fixed number  $k$  of nearest electrodes, or by using a search radius  $d$  for neighbouring electrodes. In the latter case, the number of  $p$ -values used for correction may differ across electrodes and will depend on the local density of the montage used –  $p$ -values associated with electrodes in densely populated regions of the electrode cap will be corrected using larger sets of  $p$ -values than those associated with electrodes of more sparsely populated areas.

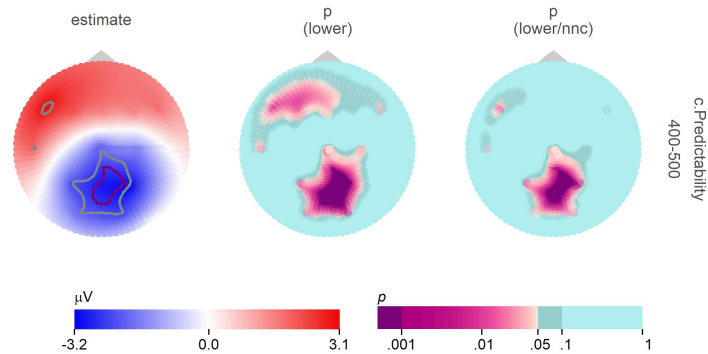
In the next step, the example ERP data for two 100 ms time windows covering the N400 time range are modelled using the same model definition as for the `wfmm`-based model (section 8.6.2.4 above). While the model formula is identical, the model fitted with `lmerMom` will incorporate correlations between the random by-subject intercepts and by-subject random slopes for the 'c.Predictability' effect. The `nnc` argument sets a search radius of (approximately) five cm for the nearest-neighbour  $p$ -value correction.

```
n400.avg.lmermom <-  
  lmerMom(x = n400.avg.base,  
    formula = ~ 1 + c.Session * c.Predictability +  
      (1 + c.Predictability | Subject),  
    time.windows = seq(300, 500, by = 100),  
    nnc = list(correct = TRUE, d = 5))
```

Once fitted, the `mom` function is applied to the resulting `lmerMom` object to summarise either the fixed effects or variance components for the time windows used for analysis. Figure 8.12 summarises the `n400.avg.lmermom` object as model maps, plotting the estimate along with  $p$ -values based upon lower-bound degrees of freedom and their nearest-neighbour corrected analogues for the later 400–500 ms time window. The contours in the estimate maps demarcate significant regions as flagged by the nearest-neighbour corrected  $p$ -values.

```
mom(n400.avg.lmermom,  
  facet.grid = time.windows + predictor ~ statistic,  
  ffx.settings = list(  
    predictors = c("c.Predictability"),  
    statistics = c("estimate", "lower.p", "lower.p.nnc")),
```

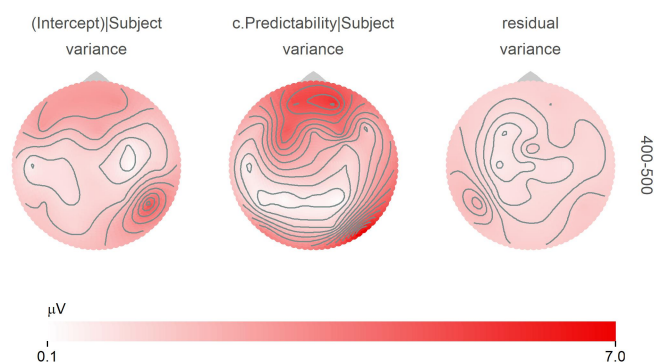
```
time.windows = c(400, 500),
contours = list(project.p = "lower.p.nnc"),
legend = list(position = "bottom"))
```



**Figure 8.12:** Model maps of the `lmerMom`-based estimates for the 'c.Predictability' fixed effect for the 400-500 ms time window.

The last call illustrates how to plot model maps for the respective variance estimates. Figure 8.13 presents the maps for the by-subject random intercepts and random slopes for 'c.Predictability' as well as the residual variance.

```
mom(n400.avg.lmermom,
type = "VC",
facet.grid = time.windows ~ VC + statistic,
time.windows = c(400, 500),
legend = list(position = "bottom"))
```



**Figure 8.13:** Model maps for the variance estimates of the `lmerMom`-based model.

## 8.7 Summary

The *stepmom* package for R (R Development Core Team, 2013) comprises numerous functions for mixed model analysis of multi-channel ERP data. The package functionality includes fitting WFMMs to multi-channel ERP data without specification of temporal windows or spatial regions of interest (*wfmm*), complemented by fitting standard LMMs to time-window averages (*lmerMom*). Both of these model-fitting functions make use of the mixed model formula interface known from *lme4* (Bates et al., 2014).

As with the *wrapfmm* package, *stepmom* provides functions which allow users to flexibly explore and plot both the model output as well as the underlying data at all stages of analysis using efficient, high-level user interfaces. The *vom* and *mom* functions provide means to plot the ERP data and the modelling results on the montage-level as high quality scalp maps. The *ani* function goes one step further by producing fully controllable animations of data or model scalp maps over time, ready for insertion into HTML pages. Jointly, these tools allow to convey the rich information contained within ERP data and the *wfmm*-based models to readers with a minimal amount of loss.

In the course of the chapter, I pointed out a number of current limitations, the most serious on the modelling level certainly being the complete neglect of the spatial (and thus also the spatiotemporal) correlations inherent to multi-channel ERP data. At the moment, *stepmom*'s *wfmm* method models the single electrodes' data independently and serially. Other practical limitations include restrictions on the range of electrode montages *stepmom* currently supports as well as limits on model complexity and size of the data sets (e.g. the number of single trials or electrodes included) imposed by available computational resources. The concluding chapter discusses these issues with respect to the criterion of 'holism' as defined in chapter 6 as well as possible ways to better account for complex ERP data.



## Chapter 9

# General discussion

### 9.1 Summary

In the course of this second, methodological part of this thesis, I have reviewed a number of issues arising in the analysis of scalar and functional psycholinguistic data and have presented recent alternatives to more traditional analysis approaches, including LMMs for scalar and curve data as well as the framework of functional data analysis for general curve analysis and the WFMM of Morris and Carroll (2006) as a specific functional analysis tool. I then surveyed two packages for the R environment: *wrapfmm* works as an R interface for the WFMM software and the *stepmom* package allows WFMM-based analysis of multichannel ERP-data.

The discussion in chapter 6 showed that a number of additional challenges come into play when the data to be analysed are curves, rather than scalar 'point' data such as reaction times: the ideal analysis tool should again be able to make use of all the information contained in these rich data types, which is not only information related to the design structure, but also curve-internal information such as the correlation between adjacent and nearby grid points, to allow for appropriate estimation of the exact onset, offset, extent and shape of effects of interest. The latter is part of an additional layer of relevant structure I have referred to as object structure, which is determined by the biological objects and processes generating the signals analysed. I suggested to label models which allow to capture the detailed dynamics of effects within curves and are able to account for relevant aspects of design and object structure as holistic, since they make use of all the information available with complex curve data. Ideally, the analysis of curve data which goes beyond prototypical one-dimensional curves, such as spatiotemporal ERP data, should also be done in a maximally holistic way, acknowledging not only the presence of a temporal object axis, but also the spatial as well as spatiotemporal dimensions.

A brief survey of standard approaches to curve- and ERP-analysis showed that canonical analyses mostly do not make use of the rich information contained in these data sets in the

best way possible – rather, they often maximise the amount of information *loss*: not only do they involve averaging within subjects (and/or items), but also feature extraction along all of the object axes. In standard ERP analyses, for example, the data are aggregated across time within temporal windows and pooled across space (electrodes) within spatial regions of interest. This coercion of curve data to points precludes obtaining detailed information about the extent and dynamics of effects across time (and/or space) and entails the risk of missing relevant effects when the a-priori parameters for windows/regions of interest are ill chosen or when some predictors show unexpected – but potentially relevant – effects outside standard intervals and/or regions.

The *wrapfmm* and *stepmom* packages introduced here were developed to provide low-barrier access to the WFMM, a model which aims at the holistic analysis of functional data. Both packages face a number of limitations, however: for example, the *wfmm* method of the *stepmom* package performs the modelling of ERP data separately for each electrode, neglecting the spatial and spatiotemporal object structure inherent to ERP data. The following section takes a closer look at some of the most practically relevant limitations, discusses their implications, strategies for coping with these as well as possible extensions.

## 9.2 Discussion

### 9.2.1 Meeting and failing holism with the current approach in practice

Starting at the design structure level, the WFMM covers almost the same range of features available with, for example, *lme4*'s `lmer` function: it allows to model the effects of factorial and continuous predictors at the single-trial level and accommodates multiple levels of random effects in nested or crossed configurations. An option available with `lmer`, but missing in the current implementation of the WFMM is the inclusion of random correlation parameters modelling correlation between random effects such as, for example, by-subject random intercepts and slopes. While Barr et al. (2013) strongly recommend the use of maximal random effects structures with LMMs in order to keep the Type I error rate at the nominal  $\alpha$ -level, they also note that models with complex random effects specifications frequently fail to converge. In their discussion of different coping strategies, they point out that models missing random correlation parameters performed similar to LMMs with maximal random effects structures and recommend the removal of correlation parameters as one of the primary strategies to simplify the random effects structure of non-converging models. While the lack of random correlations is a limiting factor on the way to maximal WFMMs, its consequences with respect to generalisability of the results may thus not be too serious.

In practice, a more problematic factor may be the number of fixed and – especially – random effects included in a model, since the computing time scales linearly with the number of grid points (i.e.  $T$ ), but scales less well with the number of fixed and random effects. While the WFMM can in principle handle (crossed) random effects for subjects and items within one model, it may be infeasible to specify (near) maximal random effects structures in some

cases. Such cases may arise with experimental designs requiring multiple (by-subject and/or by-item) random slopes and/or with experiments making use of a high number of items and a within-item design, where really complex models may take a substantial amount of time to finish. One possibility to make the fitting of complex models more feasible would be to implement a user-friendly parallelization strategy, e.g. for the MCMC performed.

When modelling multi-channel ERP data, the time factor arising with complex models becomes even more relevant. Thus, depending on the experimental design and the computational resources available, one may have to compromise on the random effects structure for the sake of computational feasibility. To compensate somewhat for these practical limitations, the *stepmom* package offers the `lmerMom` function to allow comparisons of reduced WFMMs with the output of (near) maximal standard LMMs fitted to time-window average data.

Where possible, the use of experimental designs requiring, for example, no random effects for items may already significantly reduce the computational burden. However, this will often not be a viable alternative.

Similar issues arise with memory resources and large data sets may exceed the available RAM on standard computers. In fact, in big data settings, even loading the entire data into memory at once may become impossible. In such cases, using a computer with ample of memory may of course be a solution, but again it will not be an option for everybody. A more general remedy would be the implementation of sequential methods for model fitting within the core executable, which do not require access to the full data set at once. Crainiceanu et al. (2013) discuss computational challenges arising with multilevel functional (big) data and Herrick and Morris (2006) provide more information on computational aspects of the WFMM and formulas for estimating disk and RAM usage for a WFMM.

On the object level, *wrapfmm* as presented here falls short of supporting the multidimensional 2d-image and 3d-volume modelling implemented in the WFMM executable. For some users this will be a significant shortcoming, since multidimensional functional data frequently occur with psychology and psycholinguistics – provided that computational resources are not an issue, time-frequency images derived from ERPs or different kinds of phonetic data sets (e.g. Lancia et al., 2015) may be modelled holistically as two-dimensional functional data. Similarly, structural brain scans may be modelled using 3d-WFMMs.

One of the challenges coming with multidimensional data is that they are often high-dimensional (even when sampled at a relatively coarse grid) and can easily get huge, especially when tackled at the single-trial level. Implementing support for multidimensional WFMMs may thus require to add functionalities for working with big data not only in the core software, but also within R using one of the available packages allowing the handling and manipulation of out-of-memory data sets (e.g. Adler, Gläser, Nenadic, Oehlschlägel, & Zucchini, 2013; Wickham & Francois, 2014).

As already pointed out, by performing the modelling separately for each electrode, the *stepmom* package currently falls short of taking into account the spatial (and thus also the spatiotemporal) axes of the object structure during modelling, which is inherent to multichannel

EEG data. Consequently, it only goes halfway towards an adequate exploitation of the the full object structure and towards fully holistic models, similar to other approaches to EEG-data analysis outlined in section 6.4.2, including the point-by-point/mass univariate type (e.g. Amsel, 2011b; Hauk et al., 2012, 2006, 2009; Pernet, Chauveau, et al., 2011; Smith & Kutas, 2014b) and electrode-wise GAM-based models (Kryuchkova et al., 2012; Tremblay & Baayen, 2010).

Rousselet and Pernet (2011) stress the importance of using robust statistics with modern EEG analysis methods which try to capture the detailed spatiotemporal dynamics of effects, including the application of appropriate statistical models and measures which are robust against deviations from the assumed distribution parameters as well as adequate procedures for multiplicity correction. According to Rousselet and Pernet, the latter should not revert to ad hoc criteria (such as dismissing all effects which last less than a fixed number of consecutive time points), but should be data-driven and take into account the inherent correlation structure.

The MULTIBONDS and SIMBAS (Meyer et al., 2015) presented in section 7.7.2 should meet these criteria, since they are a natural by-product of the modelling and solely rely on the posterior samples obtained. However, when making inferences about multichannel ERP data, they only represent corrections over time, but the spatial axis is currently neglected. In principle, both of these multiplicity corrected probability types can be computed for posteriors for multi-dimensional models in a straightforward manner, but they may not be the most efficient correction for really high-dimensional data. A more detailed investigation and comparison with other multiplicity corrections procedures under different conditions may reveal their strengths and weaknesses and lead to adaptations for different kinds of functional data.

Zhu et al. (2011) present a robust version of the WFMM (R-WFMM) which has the effect of down-weighting outlying curves – or outlying regions within single curves – for estimation. The R-WFMM thus identifies and removes spurious patterns within functions, but retains their true features, which can lead to greatly enhanced precision in the estimates and improved adaptive regularisation. Since outliers are almost always present in complex functional data sets, even with care in the preprocessing and correction and removal of artefacts, the availability of a robust modelling option within the core WFMM software would be a valuable extension for many applications.

## **9.2.2 Possible extensions for modelling EEG data with the WFMM**

The preceding section already listed a number of desirable WFMM-features and possible topics for further research. In the following, I will brainstorm briefly on further possible extensions of the WFMM and its interface package with a focus on modelling of EEG data.

So far, I have deliberately left another central object-structural axis of EEG data out of the discussion: the frequency dimension, which holds additional information at the object level (see Makeig, Debener, Onton, & Delorme, 2004, for example). As outlined in section 6.6.2, the DWT decomposes the original signals at multiple scales representing different

frequency ranges (also see Raz, Dickerson, & Turetsky, 1999); however, the standard DWT only provides little flexibility in determining frequency ranges of interest.

One way to enhance flexibility here is presented by Samar et al. (1999) and Raz et al. (1999), who discuss the value of a specific variant of the DWT for the analysis of ERP data: the use of 'wavelet packets' allows to split the separate scales returned by a discrete wavelet transform into further subscales, giving the user more fine control over the frequency ranges covered by the wavelet scales. If the use of the DWT with wavelet packets is a viable option within the WFMM framework, modelling and inference in the time-frequency domain could be performed with a higher degree of flexibility.

Further, Duru, Ademoglu, and Demiralp (2009) present an interesting application of the wavelet transform in the *spatial* domain to enhance reconstruction of EEG sources: in their approach, the electrodes are first projected onto a rectangular grid using an algorithm which starts from a realistic head model and minimises spatial distortions during projection. After transforming the ERP data of each electrode to the time-frequency domain using the 1d-DWT, the spatial topographic maps of selected wavelet coefficients can be further decomposed to spatial subtopographic maps representing different spatial frequency extents using the 2d-DWT. These spatial submaps, in turn, can be fed to source reconstruction algorithms for improved source localisation of the underlying neural generators (see also Samar et al., 1999).

With a fully holistic WFMM-based modelling of ERP data which also accounts for the spatial dimension, such spatial submaps for the various model parameters could be a product of the fitting process. With a projection of the electrodes onto a 2d-grid and subsequent interpolation along the lines of Duru et al. at each time point, the existing support for 3d-WFMMs may actually be used to model ERPs as three-dimensional event-related spatiotemporal volumes (see also Litvak et al., 2011). The output of such WFMMs would include posterior samples for each wavelet scale and may be used for inference in the reconstructed data space as well as at specific wavelet scales for flexible inference on the various object axes.

When operating at the single-trial level, such a holistic multidimensional modelling may become a considerable computational challenge in terms of memory resources and computation time, even with relatively coarse-grained resolutions in all dimensions, since the number of trials in ERP studies can easily range into the thousands. Possible strategies may include the use of wavelet compression as well as applying sequential model fitting procedures and alternatives to MCMC-based estimation of the posterior, such as (online) variational Bayes algorithms (e.g. Kabisa, Dunson, & Morris, 2015). Alternative ways of providing for exhaustively holistic (W)FMMs may involve some new strategy involving a combination of basis functions, decompositions and transformations to make most out of the data yet ensure computational feasibility (J.S. Morris, personal communication). In recent work, Zhu et al. (accepted) present an analysis of ERP data within the (W)FMM framework which takes spatial correlation between electrodes into account by assuming a Matérn structure, thus providing

one possible and promising approach to still more holistic modelling.

### 9.2.3 Improving information transmission

The core aim of holistic analysis approaches is to maintain and exploit the information present in the data to the maximum degree possible. As we move closer to this goal, we also need to make sure that researchers also get access to tools which produce summaries optimised for the respective data type and allow to communicate the analysis results to others in a flexible manner which maximises information *transmission*.

Ideally, this will involve the possibility for others to interact with the underlying data, the model results and the output of the summary functions with as few barriers as possible. In many cases, summaries of complex, possibly multi- or high-dimensional functional data and models in research reports will focus on those aspects which are of interest for the original question of the study. Thus, in a report on an ERP experiment involving two or three predictors of interest and a (potentially larger) number of control covariates, the authors will necessarily concentrate on summarising and illustrating the relevant effects of their core variables.

However, what is a mere nuisance variable which has to be controlled for for one researcher (say, word frequency or length), may be the focus of interest for another one. Rather than losing all that information in the course of the reporting process, it should be made available for further mining. With the *stepmom* package, for example, posterior estimate and probability maps (as well as the underlying data) can be rendered as animations over time, conveying the full spatiotemporal information contained within the models. Such animated model maps can be posted for all model predictors as supplementary materials and readers can explore the results for any variable and any time period interactively. Such interactive graphical model summaries not only ensure maintenance of information about the effects of control covariates not presented in the original reports, but also about effects of main predictors falling outside of time windows reported upon.

Interactivity with the data and model output in the form of the WFMM's posteriors or the summary objects can of course be taken to another level by making either available for public download. In the case of the data and posteriors, this may become a challenging task, depending on the size of either. A WFMM's posteriors can quickly exceed the size of the original data by far and for high- and multidimensional functional data sets, it may become too big for straightforward up- and download. However, even in such cases the summary objects are compact enough in size to be posted easily as supplementary material and readers can use these to investigate any aspects of interest on their own.

## 9.3 Conclusion

Functional data are increasingly encountered in experimental studies and hold rich information about the temporal, spatial or spatiotemporal extent and course of effects of interest – yet, much of this information is lost with standard analysis paradigms, which involve coercion of

functional data to point data via feature extraction on the object level as well as averaging on the design level. Functional mixed effects models such as the WFMM (Morris & Carroll, 2006) offer an alternative for analysing different kinds of curve data making maximal use of the information present at both levels, acting as (nearly) holistic modelling tools.

I presented two R packages that allow user-friendly usage of the WFMM for general curve- and ERP-analysis, pointed out a number of limitations of the packages as introduced here and discussed a number of additional features and extensions both packages as well as the core WFMM software would profit from. As stressed at several points in this thesis, the practical impact of the current package versions is limited, but the new packages that are currently being prepared for release at the time of publication will already address some of the limitations discussed and implement numerous additional features (see the preface for more information).





# Bibliography

- Abbott, B. (2003). Montague grammar. In W. J. Frawley (Ed.), *International encyclopedia of linguistics* (Second ed., Vol. 3, pp. 79–81). New York: Oxford University Press.
- Ackerman, F., & Moore, J. (2001). *Proto-properties and grammatical encoding*. Stanford: CSLI Publications.
- Adler, D., Gläser, C., Nenadic, O., Oehlschlägel, J., & Zucchini, W. (2013). ff: memory-efficient storage of large data on disk and fast access functions [Computer software manual]. Retrieved from <http://CRAN.R-project.org/package=ff> (R package version 2.2-12)
- Alama, J. (2015). The Lambda calculus. In E. N. Zalta (Ed.), *The stanford encyclopedia of philosophy* (Spring 2015 ed.). Retrieved from <http://plato.stanford.edu/archives/spr2015/entries/lambda-calculus/>
- Alexiadou, A. (2010). Nominalizations: A probe into the architecture of grammar – part I: The nominalization puzzle. *Language and Linguistics Compass*, 4(7), 496–511.
- Amsel, B. D. (2011a). *Semantic feature type constrains the organization and computation of concrete conceptual knowledge* (PhD thesis). University of Toronto.
- Amsel, B. D. (2011b). Tracking real-time neural activation of conceptual knowledge using single-trial event-related potentials. *Neuropsychologia*, 49(5), 970–983.
- Anderson, M. (1979). *Noun phrase structure* (PhD thesis). The University of Connecticut.
- Anderson, M. (2006). Affectedness. In M. Everaert & H. van Riemsdijk (Eds.), *The blackwell companion to syntax* (Vol. I, pp. 121–141). Malden: Blackwell.
- Aston, J. A., Chiou, J.-M., & Evans, J. P. (2010). Linguistic pitch analysis using functional principal component mixed effect models. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 59(2), 297–317.
- Azizian, A., Freitas, A. L., Parvaz, M. A., & Squires, N. K. (2006). Beware misleading cues: Perceptual similarity modulates the N2/P3 complex. *Psychophysiology*, 43(3), 253–260.
- Baayen, R. H. (2010). A real experiment is a factorial experiment? *The Mental Lexicon*, 5(1), 149–157.
- Baayen, R. H. (2013). languageR: Data sets and functions with "Analyzing Linguistic Data: A practical introduction to statistics". [Computer software manual]. Retrieved from <http://CRAN.R-project.org/package=languageR> (R package version 1.4.1)
- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed

- random effects for subjects and items. *Journal of Memory and Language*, 59 (4), 390–412.
- Baayen, R. H., & Milin, P. (2010). Analyzing reaction times. *International Journal of Psychological Research*, 3 (2), 12–28.
- Baayen, R. H., Piepenbrock, R., & van Rijn, H. (1993). *The CELEX lexical data base on CD-ROM*. Philadelphia: Linguistic Data Consortium.
- Baggio, G., Choma, T., van Lambalgen, M., & Hagoort, P. (2010). Coercion and compositionality. *Journal of Cognitive Neuroscience*, 22 (9), 2131–2140.
- Bagiella, E., Sloan, R. P., & Heitjan, D. F. (2000). Mixed-effects models in psychophysiology. *Psychophysiology*, 37 (1), 13–20.
- Baker, M. (1996). On the structural positions of themes and goals. In J. Rooryck & L. Zaring (Eds.), *Phrase structure and the lexicon* (pp. 7–34). Dordrecht: Kluwer.
- Bandecchi, V., & Keane, M. T. (2013). Lexical decomposition in understanding Italian and English motion verb. *Contemporary Linguistics (Suvremena lingvistika)* (76), 107–125.
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68 (3), 255–278.
- Bates, D. M. (2010). *lme4: Mixed-effects modeling with R*. Retrieved from <http://lme4.r-forge.r-project.org/lmmwR/>
- Bates, D. M., Maechler, M., & Bolker, B. (2012). *lme4: Linear mixed-effects models using S4 classes* [Computer software manual].
- Bates, D. M., Maechler, M., Bolker, B., & Walker, S. (2014). *lme4: Linear mixed-effects models using Eigen and S4* [Computer software manual]. Retrieved from <http://CRAN.R-project.org/package=lme4> (R package version 1.1-7)
- Batiukova, O., Bertinetto, P. M., Lenci, A., & Zarccone, A. (2014). Identifying actional features through semantic priming: Cross-Romance comparison. *Quaderni del Laboratorio di Linguistica*, 13, 1–31.
- Beavers, J. (2010). The structure of lexical meaning: Why semantics really matters. *Language*, 86 (4), 821–864.
- Beavers, J. (2011). On affectedness. *Natural Language & Linguistic Theory*, 29 (2), 335–370.
- Beavers, J. (2013). Aspectual classes and scales of change. *Linguistics*, 51 (4), 681–706.
- Bedny, M., Dravida, S., & Saxe, R. (2014). Shindigs, brunches, and rodeos: The neural basis of event words. *Cognitive, Affective, & Behavioral Neuroscience*, 14 (3), 891–901.
- Belletti, A., & Rizzi, L. (1988). Psych-verbs and  $\theta$ -theory. *Natural Language & Linguistic Theory*, 6 (3), 291–352.
- Bengtsson, H., & Riedy, J. (2011). *R.matlab: Read and write of MAT files together with R-to-matlab connectivity* [Computer software manual]. Retrieved from <http://CRAN.R-project.org/package=R.matlab> (R package version 1.5.2)
- Berkum, J. v., Hagoort, P., & Brown, C. (1999). Semantic integration in sentences and discourse: Evidence from the N400. *Journal of Cognitive Neuroscience*, 11 (6), 657–671.

- Bierwisch, M. (1989). Event nominalization: Proposals and problems. In W. Motsch (Ed.), *Wortstruktur und Satzstruktur (Linguistische Studien 194)* (pp. 1–73). Berlin: Akademie der Wissenschaften der DDR; Zentralinstitut für Sprachwissenschaft.
- Bierwisch, M. (2009). Nominalization – lexical and syntactic aspects. In A. Giannakidou & M. Rathert (Eds.), *Quantification, definiteness, and nominalization* (pp. 281–320). Oxford: Oxford University Press.
- Bivand, R. (2013). spdep: Spatial dependence: Weighting schemes, statistics and models [Computer software manual]. Retrieved from <http://CRAN.R-project.org/package=spdep> (R package version 0.5-68)
- Blume, K. (1998). A contrastive analysis of interaction verbs with dative complements. *Linguistics*, 36 (2), 253–280.
- Blume, K. (2000). *Markierte Verbvalenzen im Sprachvergleich: Lizenzierungs- und Linkingbedingungen*. Tübingen: Niemeyer.
- Boland, J. E. (1997). The relationship between syntactic and semantic processes in sentence comprehension. *Language and Cognitive Processes*, 12 (4), 423–484.
- Bonnotte, I. (2008). The role of semantic features in verb processing. *Journal of Psycholinguistic Research*, 37 (3), 199–217.
- Bornkessel, I., McElree, B., Schlesewsky, M., & Friederici, A. D. (2004). Multi-dimensional contributions to garden path strength: Dissociating phrase structure from case marking. *Journal of Memory and Language*, 51 (4), 495–522.
- Bornkessel, I., & Schlesewsky, M. (2006). The extended argument dependency model: A neurocognitive approach to sentence comprehension across languages. *Psychological Review*, 113 (4), 787.
- Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2008). An alternative perspective on semantic P600 effects in language comprehension. *Brain Research Reviews*, 59 (1), 55–73.
- Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2009). The role of prominence information in the real-time comprehension of transitive constructions: A cross-linguistic approach. *Language and Linguistics Compass*, 3 (1), 19–58.
- Brennan, J., & Pykkänen, L. (2008). Processing events: Behavioral and neuromagnetic correlates of aspectual coercion. *Brain and Language*, 106 (2), 132–143.
- Brennan, J., & Pykkänen, L. (2010). Processing psych verbs: Behavioural and MEG measures of two different types of semantic complexity. *Language and Cognitive Processes*, 25 (6), 777–807.
- Bresnan, J., & Kanerva, J. M. (1989). Locative inversion in Chicheŵa: A case study of factorization in grammar. *Linguistic Inquiry*, 20 (1), 1–50.
- Brown, C., & Hagoort, P. (1993). The processing nature of the N400: Evidence from masked priming. *Journal of Cognitive Neuroscience*, 5 (1), 34–44.
- Bücking, S. (2010). Zur Interpretation adnominaler Genitive bei nominalisierten Infinitiven im Deutschen. *Zeitschrift für Sprachwissenschaft*, 29 (1), 39–77.
- Bücking, S. (2011). *Kompositional flexibel – Partizipanten und Modifikatoren in der Nominaldomäne* (PhD thesis). Universität Tübingen.

- Bugli, C., & Lambert, P. (2006). Functional ANOVA with random functional effects: An application to event-related potentials modelling for electroencephalograms analysis. *Statistics in Medicine*, 25(21), 3718–3739.
- Burkhardt, P. (2006). Inferential bridging relations reveal distinct neural mechanisms: Evidence from event-related brain potentials. *Brain and Language*, 98(2), 159–168.
- Burkhardt, P. (2007). The P600 reflects cost of new information in discourse memory. *Neuroreport*, 18(17), 1851–1854.
- Burns, M. D., Bigdely-Shamlo, N., Smith, N. J., Kreutz-Delgado, K., & Makeig, S. (2013). Comparison of averaging and regression techniques for estimating event related potentials. In *Engineering in medicine and biology society (EMBC), 35th annual international conference of the IEEE* (pp. 1680–83).
- Carreiras, M., Salillas, E., & Barber, H. (2004). Event-related potentials elicited during parsing of ambiguous relative clauses in Spanish. *Cognitive Brain Research*, 20(1), 98–105.
- Chomsky, N. (1970). Remarks on nominalization. In R. A. Jacobs & P. S. Rosenbaum (Eds.), *Readings in English transformational grammar* (pp. 184–221). Waltham: Ginn.
- Choudhary, K. K., Schlesewsky, M., Roehm, D., & Bornkessel-Schlesewsky, I. (2009). The N400 as a correlate of interpretively relevant linguistic rules: Evidence from Hindi. *Neuropsychologia*, 47(13), 3012–3022.
- Christensen, R. H. B. (2013a). *Analysis of ordinal data with cumulative link models – estimation with the R-package ordinal*. Retrieved from [http://cran.r-project.org/web/packages/ordinal/vignettes/clm\\_intro.pdf](http://cran.r-project.org/web/packages/ordinal/vignettes/clm_intro.pdf)
- Christensen, R. H. B. (2013b). *ordinal – regression models for ordinal data* [Computer software manual]. Retrieved from <http://www.cran.r-project.org/package=ordinal/> (R package version 2013.9-30)
- Church, A. (1936). An unsolvable problem of elementary number theory. *American Journal of Mathematics*, 58(2), 345–363.
- Clark, H. H. (1973). The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior*, 12(4), 335–359.
- Cohen, J. (1983). The cost of dichotomization. *Applied Psychological Measurement*, 7(3), 249–253.
- Coleman, E. B. (1964). Generalizing to a language population. *Psychological Reports*, 14(1), 219–226.
- Coll-Florit, M., & Gennari, S. P. (2011). Time in language: Event duration in language comprehension. *Cognitive Psychology*, 62(1), 41–79.
- Comrie, B. (1989). *Language universals and linguistic typology* (Second ed.). Chicago: University of Chicago Press.
- Coulson, S., King, J. W., & Kutas, M. (1998). Expect the unexpected: Event-related brain response to morphosyntactic violations. *Language and Cognitive Processes*, 13(1), 21–58.
- Crainiceanu, C. M., Caffo, B. S., & Morris, J. S. (2013). Multilevel functional data analysis.

- In J. S. Simonoff, M. A. Scott, & B. D. Marx (Eds.), *The SAGE handbook of multilevel modeling* (pp. 223–48). London: SAGE.
- Crainiceanu, C. M., Ruppert, D., Carroll, R., Joshi, A., & Goodner, B. (2007). Spatially adaptive bayesian penalized splines with heteroscedastic errors. *Journal of Computational and Graphical Statistics*, 16 (2), 265–288.
- Croft, W. (1993). Case marking and the semantics of mental verbs. In J. Pustejovsky (Ed.), *Semantics and the lexicon* (pp. 55–72). Dordrecht: Kluwer.
- Croft, W. (1998). Event structure in argument linking. In M. Butt & W. Geuder (Eds.), *The projection of arguments: Lexical and compositional factors* (pp. 21–63). Stanford: CSLI Publications.
- Cruse, D. A. (1973). Some thoughts on agentivity. *Journal of Linguistics*, 9(1), 11–23.
- Cutler, A. (1981). Making up materials is a confounded nuisance, or: Will we able to run any psycholinguistic experiments at all in 1990? *Cognition*, 10(1), 65–70.
- Czypionka, A. (2014). *The interplay of object animacy and verb class in representation building* (PhD Thesis). Humboldt-Universität zu Berlin.
- Dambacher, M., Kliegl, R., Hofmann, M., & Jacobs, A. M. (2006). Frequency and predictability effects on event-related potentials during reading. *Brain Research*, 1084 (1), 89–103.
- Dambacher, M., Rolfs, M., Göllner, K., Kliegl, R., & Jacobs, A. M. (2009). Event-related potentials reveal rapid verification of predicted visual input. *PLOS ONE*, 4 (3), e5047.
- Davidson, D. J. (2009). Functional mixed-effect models for electrophysiological responses. *Neurophysiology*, 41 (1), 71–79.
- Davidson, D. J., Hanulíková, A., & Indefrey, P. (2012). Electrophysiological correlates of morphosyntactic integration in German phrasal context. *Language and Cognitive Processes*, 27(2), 288–311.
- Davis, A. R. (2011). Thematic roles. In C. Maienborn, K. von Stechow, & P. Portner (Eds.), *Semantics: An international handbook of natural language meaning* (Vol. 1, pp. 399–420). Berlin: de Gruyter.
- Davis, A. R., & Koenig, J. P. (2000). Linking as constraints on word classes in a hierarchical lexicon. *Language*, 76 (1), 56–91.
- de Almeida, R. G. (2004). The effect of context on the processing of type-shifting verbs. *Brain and Language*, 90 (1), 249–261.
- de Almeida, R. G., & Dwivedi, V. D. (2008). Coercion without lexical decomposition: Type-shifting effects revisited. *Canadian Journal of Linguistics*, 53 (2/3), 301–326.
- DeLong, K. A., Quante, L., & Kutas, M. (2014). Predictability, plausibility, and two late ERP positivities during written sentence comprehension. *Neuropsychologia*, 61, 150–162.
- DeLong, K. A., Urbach, T. P., Groppe, D. M., & Kutas, M. (2011). Overlapping dual ERP responses to low cloze probability sentence continuations. *Psychophysiology*, 48 (9), 1203–1207.
- De Wit, P., & Schoorlemmer, M. (1996). Prenominal arguments in Russian, German and Dutch. *ZAS Papers in Linguistics*, 5, 184–202.
- Dien, J., & Santuzzi, A. M. (2005). Application of repeated measures ANOVA to high-density ERP datasets: A review and tutorial. In T. C. Handy (Ed.), *Event-related potentials: A*

- methods handbook*. Cambridge, MA: MIT Press.
- Dowty, D. (1979). *Word meaning and Montague grammar*. Dordrecht: Reidel.
- Dowty, D. (1991). Thematic proto-roles and argument selection. *Language*, 67, 547–619.
- Duru, A. D., Ademoglu, A., & Demiralp, T. (2009). Analysis of brain electrical topography by spatio-temporal wavelet decomposition. *Mathematical and Computer Modelling*, 49(11), 2224–2235.
- Ehrich, V. (2002). On the verbal nature of certain nominal entities. In I. Kaufmann & B. Stiebels (Eds.), *More than verbs. A Festschrift for Dieter Wunderlich* (pp. 69–89). Berlin: Akademie Verlag.
- Ehrich, V. (2003). The thematic interpretation of plural nominalizations. In H. Härtl & H. Tappe (Eds.), *Mediating between concepts and grammar* (pp. 351–378). Berlin: De Gruyter.
- Ehrich, V., & Rapp, I. (2000). Sortale Bedeutung und Argumentstruktur: *ung-* Nominalisierungen im Deutschen. *Zeitschrift für Sprachwissenschaft*, 19(2), 245–303.
- Ettinger, A., Linzen, T., & Marantz, A. (2014). The role of morphology in phoneme prediction: Evidence from MEG. *Brain and Language*, 129, 14–23.
- Fazio, M. A., Grytz, R., Morris, J. S., Bruno, L., Gardiner, S. K., Girkin, C. A., & Downs, J. C. (2014). Age-related changes in human peripapillary scleral strain. *Biomechanics and Modeling in Mechanobiology*, 13(3), 551–563.
- Federmeier, K. D., Wlotko, E. W., De Ochoa-Dewald, E., & Kutas, M. (2007). Multiple effects of sentential constraint on word processing. *Brain Research*, 1146, 75–84.
- Ferretti, T. R., McRae, K., & Hatherell, A. (2001). Integrating verbs, situation schemas, and thematic role concepts. *Journal of Memory and Language*, 44(4), 516–547.
- Filip, H. (2011). Aspectual class and Aktionsart. In C. Maienborn, K. von Heusinger, & P. Portner (Eds.), *Semantics: An international handbook of natural language meaning* (Vol. 2, pp. 1186–1217). Berlin: de Gruyter.
- Fillmore, C. J. (1968). The Case for Case. In E. Bach & R. T. Harms (Eds.), *Universals in linguistic theory* (pp. 1–88). New York: Holt, Rinehart and Winston.
- Fillmore, C. J. (1970). The grammar of 'hitting' and 'breaking'. In R. Jacobs & P. S. Rosenbaum (Eds.), *Readings in English transformational grammar* (pp. 120–33). Waltham: Ginn.
- Fillmore, C. J. (1977). The case for case reopened. In P. Cole & J. Sadock (Eds.), *Grammatical relations: Syntax and semantics 8* (pp. 59–82). New York: Academic Press.
- Fodor, J. A., Garrett, M. F., Walker, E. C., & Parkes, C. H. (1980). Against definitions. *Cognition*, 8(3), 263–367.
- Folstein, J. R., & Van Petten, C. (2004). Multidimensional rule, unidimensional rule, and similarity strategies in categorization: Event-related brain potential correlates. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30(5), 1026.
- Folstein, J. R., & Van Petten, C. (2011). After the P3: Late executive processes in stimulus categorization. *Psychophysiology*, 48(6), 825–841.
- Forster, K. I., & Dickinson, R. (1976). More on the language-as-fixed-effect fallacy: Monte carlo estimates of error rates for  $F_1$ ,  $F_2$ ,  $F'$ , and  $\min F'$ . *Journal of Verbal Learning and*

- Verbal Behavior*, 15(2), 135–142.
- Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology*, 14(2), 178–210.
- Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences*, 6(2), 78–84.
- Friederici, A. D., Hahne, A., & Mecklinger, A. (1996). Temporal structure of syntactic parsing: early and late event-related brain potential effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22(5), 1219.
- Friederici, A. D., Hahne, A., & Saddy, D. (2002). Distinct neurophysiological patterns reflecting aspects of syntactic complexity and syntactic repair. *Journal of Psycholinguistic Research*, 31(1), 45–63.
- Friederici, A. D., Pfeifer, E., & Hahne, A. (1993). Event-related brain potentials during natural speech processing: Effects of semantic, morphological and syntactic violations. *Cognitive Brain Research*, 1(3), 183–192.
- Frisch, S., & Schlesewsky, M. (2001). The N400 reflects problems of thematic hierarchizing. *NeuroReport*, 12(15), 3391.
- Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J.-P., Frith, C. D., & Frackowiak, R. S. (1994). Statistical parametric maps in functional imaging: A general linear approach. *Human Brain Mapping*, 2(4), 189–210.
- Fuß, E. (2011). Eigennamen und adnominaler Genitiv im Deutschen. *Linguistische Berichte*, 2011(225), 19–42.
- Geangu, E., Hauf, P., Bhardwaj, R., & Bentz, W. (2011). Infant pupil diameter changes in response to others' positive and negative emotions. *PLOS ONE*, 6(11), e27132.
- Gelman, A. (2008). Scaling regression inputs by dividing by two standard deviations. *Statistics in Medicine*, 27(15), 2865–2873.
- Gennari, S., & Poeppel, D. (2003). Processing correlates of lexical semantic complexity. *Cognition*, 89(1), B27–B41.
- Gleitman, L. (1990). The structural sources of verb meanings. *Language acquisition*, 1(1), 3–55.
- Gordon, P. C., Hendrick, R., & Johnson, M. (2001). Memory interference during language processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27(6), 1411.
- Grimm, S. (2011). Semantics of case. *Morphology*, 21(3), 515–544.
- Grimshaw, J. (1990). *Argument structure*. Cambridge: MIT Press.
- Grimshaw, J. (2005a). Extended projection. In *Words and structure* (pp. 1–74). Stanford: CSLI Publications. (Based on an unpublished manuscript from 1991)
- Grimshaw, J. (2005b). Semantic structure and semantic content. In *Words and structure* (pp. 75–89). Stanford: CSLI Publications. (Based on an unpublished manuscript from 1993)
- Grimshaw, J. (2011). Deverbal nominalization. In C. Maienborn, K. von Stechow, & P. Portner (Eds.), *Semantics: An international handbook of natural language meaning* (Vol. 2,

- pp. 1292–1313). Berlin: de Gruyter.
- Gropen, J., Pinker, S., Hollander, M., & Goldberg, R. (1991). Affectedness and direct objects: The role of lexical semantics in the acquisition of verb argument structure. *Cognition*, 41(1), 153–195.
- Groppe, D. M., Urbach, T. P., & Kutas, M. (2011). Mass univariate analysis of event-related brain potentials/fields I: A critical tutorial review. *Psychophysiology*, 48(12), 1711–1725.
- Guo, W. (2002). Functional mixed effects models. *Biometrics*, 58(1), 121–128.
- Guthrie, D., & Buchwald, J. S. (1991). Significance testing of difference potentials. *Psychophysiology*, 28(2), 240–244.
- Hagoort, P., Brown, C., & Groothusen, J. (1993). The syntactic positive shift (SPS) as an ERP measure of syntactic processing. *Language and Cognitive Processes*, 8(4), 439–483.
- Hagoort, P., & Brown, C. M. (2000). ERP effects of listening to speech compared to reading: The P600/SPS to syntactic violations in spoken sentences and rapid serial visual presentation. *Neuropsychologia*, 38(11), 1531–1549.
- Hagoort, P., Brown, C. M., & Osterhout, L. (1999). The neurocognition of syntactic processing. In P. Brown Colin M. & Hagoort (Ed.), *The neurocognition of language* (pp. 273–316). Oxford: Oxford University Press.
- Hagoort, P., Hald, L., Bastiaansen, M., & Petersson, K. M. (2004). Integration of word meaning and world knowledge in language comprehension. *Science*, 304(5669), 438–441.
- Haider, H. (1988). Die Struktur der deutschen Nominalphrase. *Zeitschrift für Sprachwissenschaft*, 7(1), 32–59.
- Hartmann, K., & Zimmermann, M. (2003). Syntactic and semantic adnominal genitive. In C. Maienborn (Ed.), *(A)Symmetrien – (A)Symmetries. Beiträge zu Ehren von Ewald Lang* (pp. 171–202). Tübingen: Stauffenburg.
- Hauk, O., Coutout, C., Holden, A., & Chen, Y. (2012). The time-course of single-word reading: Evidence from fast behavioral and brain responses. *NeuroImage*, 60, 1462–1477.
- Hauk, O., Davis, M. H., Ford, M., Pulvermüller, F., & Marslen-Wilson, W. D. (2006). The time course of visual word recognition as revealed by linear regression analysis of ERP data. *NeuroImage*, 30(4), 1383–1400.
- Hauk, O., Pulvermüller, F., Ford, M., Marslen-Wilson, W., & Davis, M. (2009). Can I have a quick word? Early electrophysiological manifestations of psycholinguistic processes revealed by event-related regression analysis of the EEG. *Biological Psychology*, 80(1), 64–74.
- Haupt, F. S., Schlesewsky, M., Roehm, D., Friederici, A. D., & Bornkessel Schlesewsky, I. (2008). The status of subject–object reanalyses in the language comprehension architecture. *Journal of Memory and Language*, 59(1), 54–96.
- Hay, J., Kennedy, C., & Levin, B. (1999). Scalar structure underlies telicity in "degree achievements". In *Proceedings of Semantics and Linguistic Theory (SALT) 9* (pp. 127–



- 144).
- Heister, J., Würzner, K.-M., Bubenzer, J., Pohl, E., Hanneforth, T., Geyken, A., & Kliegl, R. (2011). dlexDB – eine lexikalische Datenbank für die psychologische und linguistische Forschung. *Psychologische Rundschau*, 62(1), 10–20.
- Herrick, R. C., & Morris, J. S. (2006). Wavelet-based functional mixed model analysis: Computational considerations. In *Proceedings of the Joint Statistical Meetings* (pp. 2051–2053).
- Hindy, N. C., Altmann, G. T., Kalenik, E., & Thompson-Schill, S. L. (2012). The effect of object state-changes on event processing: Do objects compete with themselves? *The Journal of Neuroscience*, 32(17), 5795–5803.
- Hoeks, J. C., Stowe, L. A., & Doedens, G. (2004). Seeing words in context: The interaction of lexical and sentence level information during reading. *Cognitive Brain Research*, 19(1), 59–73.
- Hopf, J.-M., Bayer, J., Bader, M., & Meng, M. (1998). Event-related brain potentials and case information in syntactic ambiguities. *Journal of Cognitive Neuroscience*, 10(2), 264–280.
- Hopper, P. J., & Thompson, S. A. (1980). Transitivity in grammar and discourse. *Language*, 56, 251–299.
- Hothorn, T., Bretz, F., & Westfall, P. (2008). Simultaneous inference in general parametric models. *Biometrical Journal*, 50(3), 346–363.
- Ille, N., Berg, P., & Scherg, M. (2002). Artifact correction of the ongoing EEG using spatial filters based on artifact and brain signal topographies. *Journal of Clinical Neurophysiology*, 19(2), 113–124.
- Jackendoff, R. (1972). *Semantic interpretation in generative grammar*. Cambridge, MA: MIT Press.
- Jackendoff, R. (1987). The status of thematic relations in linguistic theory. *Linguistic Inquiry*, 18(3), 369–411.
- Jackendoff, R. (1990). *Semantic structures*. Cambridge: MIT Press.
- Jackson, I., & Sirois, S. (2009). Infant cognition: Going full factorial with pupil dilation. *Developmental Science*, 12(4), 670–679.
- Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, 59(4), 434–446.
- Jaeggli, O. A. (1986). Passive. *Linguistic Inquiry*, 17(4), 587–622.
- Kaan, E., Harris, A., Gibson, E., & Holcomb, P. (2000). The P600 as an index of syntactic integration difficulty. *Language and Cognitive Processes*, 15(2), 159–201.
- Kaan, E., & Swaab, T. (2003). Repair, revision, and complexity in syntactic analysis: An electrophysiological differentiation. *Journal of Cognitive Neuroscience*, 15(1), 98–110.
- Kabisa, S., Dunson, D. B., & Morris, J. S. (2015). Online variational bayes inference for high-dimensional correlated data. *Journal of Computational and Graphical Statistics*.
- Keller, F., Corley, M., Corley, S., Konieczny, L., & Todirascu, A. (1998). *WebExp: A Java toolbox for web-based psychological experiments – Users' guide for WebExp 2.1*

- (Tech. Rep. No. HCRC/TR-99). Human Communication Research Centre, University of Edinburgh.
- Kemmerer, D., Castillo, J. G., Talavage, T., Patterson, S., & Wiley, C. (2008). Neuroanatomical distribution of five semantic components of verbs: Evidence from fMRI. *Brain and Language*, 107(1), 16–43.
- Kemmerer, D., & Gonzalez-Castillo, J. (2010). The two-level theory of verb meaning: An approach to integrating the semantics of action with the mirror neuron system. *Brain and Language*, 112(1), 54–76.
- Kennedy, C., & Levin, B. (2008). Measure of change: The adjectival core of degree achievements. In L. McNally & C. Kennedy (Eds.), *Adjectives and adverbs: Syntax, semantics and discourse* (pp. 156–182). Oxford: Oxford University Press.
- Kiebel, S. J., & Friston, K. J. (2004). Statistical parametric mapping for event-related potentials: I. Generic considerations. *NeuroImage*, 22(2), 492–502.
- Kiefer, M., & Pulvermüller, F. (2012). Conceptual representations in mind and brain: Theoretical developments, current evidence and future directions. *Cortex*, 48(7), 805–825.
- Kim, A., & Osterhout, L. (2005). The independence of combinatory semantic processing: Evidence from event-related potentials. *Journal of Memory and Language*, 52(2), 205–225.
- King, J. W., & Kutas, M. (1995). Who did what and when? using word-and clause-level ERPs to monitor working memory usage in reading. *Journal of Cognitive Neuroscience*, 7(3), 376–395.
- Kiparsky, P. (2009). On the architecture of Pāṇini's grammar. In G. Huet, A. Kulkarni, & P. Scharf (Eds.), *Sanskrit computational linguistics* (pp. 33–94). Springer.
- Kluender, R., & Kutas, M. (1993). Bridging the gap: Evidence from ERPs on the processing of unbounded dependencies. *Journal of Cognitive Neuroscience*, 5(2), 196–214.
- Kolk, H. H., Chwilla, D. J., van Herten, M., & Oor, P. J. (2003). Structure and limited capacity in verbal working memory: A study with event-related potentials. *Brain and Language*, 85(1), 1–36.
- Krifka, M. (1989). Nominal reference, temporal constitution and quantification in event semantics. In R. Bartsch, J. van Benthem, & P. van Emde Boas (Eds.), *Semantics and contextual expressions* (pp. 75–115). Dordrecht: Foris.
- Krifka, M. (1992). Thematic relations as links between nominal reference and temporal constitution. In I. A. Sag & A. Szabolcsi (Eds.), *Lexical matters* (pp. 29–54). Stanford: CSLI.
- Krifka, M. (1998). The origins of telicity. In S. Rothstein (Ed.), *Events and grammar* (pp. 197–235). Dordrecht: Kluwer Academic.
- Krivobokova, T., Kneib, T., & Claeskens, G. (2010). Simultaneous confidence bands for penalized spline estimators. *Journal of the American Statistical Association*, 105(490), 852–863.
- Kruschke, J. K. (2010a). Bayesian data analysis. *Wiley Interdisciplinary Reviews: Cognitive Science*, 1(5), 658–676.
- Kruschke, J. K. (2010b). *Doing Bayesian data analysis: A tutorial introduction with R and*

- BUGS*. Burlington, MA: Academic Press.
- Kruschke, J. K. (2010c). What to believe: Bayesian methods for data analysis. *Trends in Cognitive Sciences*, 14 (7), 293–300.
- Kryuchkova, T., Tucker, B. V., Wurm, L. H., & Baayen, R. H. (2012). Danger and usefulness are detected early in auditory lexical processing: Evidence from electroencephalography. *Brain and Language*, 122 (2), 81–91.
- Kuperberg, G. R. (2007). Neural mechanisms of language comprehension: Challenges to syntax. *Brain Research*, 1146, 23–49.
- Kuperberg, G. R., Caplan, D., Sitnikova, T., Eddy, M., & Holcomb, P. J. (2006). Neural correlates of processing syntactic, semantic, and thematic relationships in sentences. *Language and Cognitive Processes*, 21 (5), 489–530.
- Kuperberg, G. R., Choi, A., Cohn, N., Paczynski, M., & Jackendoff, R. (2010). Electrophysiological correlates of complement coercion. *Journal of Cognitive Neuroscience*, 22 (12), 2685–2701.
- Kuperberg, G. R., Sitnikova, T., Caplan, D., & Holcomb, P. J. (2003). Electrophysiological distinctions in processing conceptual relationships within simple sentences. *Cognitive Brain Research*, 17 (1), 117–129.
- Kutas, M. (1993). In the company of other words: Electrophysiological evidence for single-word and sentence context effects. *Language and Cognitive Processes*, 8 (4), 533–572.
- Kutas, M., & Hillyard, S. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature*, 307 (5947), 161–163.
- Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, 207 (4427), 203–205.
- Kutas, M., Van Petten, C. K., & Kluender, R. (2006). Psycholinguistics electrified II (1994–2005). In M. J. Traxler & M. A. Gernsbacher (Eds.), *Handbook of psycholinguistics* (Second ed., pp. 659–724). London: Elsevier.
- Lancia, L., Avelino, H., & Voigt, D. (2013). Measuring laryngealization in running speech: Interaction with contrastive tones in Yalálag Zapotec. In *Proceedings of Interspeech 2013*.
- Lancia, L., Rausch, P., & Morris, J. S. (2015). Automatic quantitative analysis of ultrasound tongue contours via wavelet-based functional mixed models. *Journal of the Acoustical Society of America*, 137 (2), EL178.
- Larson, R. K. (1990). Double objects revisited: Reply to Jackendoff. *Linguistic Inquiry*, 21 (4), 589–632.
- Laszlo, S., & Federmeier, K. D. (2011). The N400 as a snapshot of interactive processing: Evidence from regression analyses of orthographic neighbor and lexical associate effects. *Psychophysiology*, 48 (2), 176–186.
- Lau, E. F., Phillips, C., & Poeppel, D. (2008). A cortical network for semantics: (de)constructing the N400. *Nature Reviews Neuroscience*, 9 (12), 920–933.
- Lawrence, M. A. (2013). ez: Easy analysis and visualization of factorial experiments. [Computer software manual]. Retrieved from <http://CRAN.R-project.org/package=ez>

(R package version 4.2-2)

- Legendre, G. (2007). On the typology of auxiliary selection. *Lingua*, 117(9), 1522–1540.
- Leiken, K., & Pyllkkänen, L. (2014). MEG evidence that the LIFG effect of object extraction requires similarity-based interference. *Language, Cognition and Neuroscience*, 29(3), 381–389.
- Levin, B. (1999). Objecthood: An event structure perspective. In S. Billings, J. Boyle, & A. Griffith (Eds.), *Proceedings of the Chicago Linguistic Society 35, part 1: Papers from the main session* (pp. 223–247). Chicago: Chicago Linguistic Society.
- Levin, B., & Rappaport Hovav, M. (2005). *Argument realization*. Cambridge: Cambridge University Press.
- Levin, B., & Rappaport Hovav, M. (2011). Lexical conceptual structure. In C. Maienborn, K. von Stechow, & P. Portner (Eds.), *Semantics: An international handbook of natural language meaning* (Vol. 1, pp. 420–440). Berlin: de Gruyter.
- Levin, B., & Song, G. (1997). Making sense of corpus data: a case study of verbs of sound. *International Journal of Corpus Linguistics*, 2(1), 23–64.
- Levitin, D. J., Nuzzo, R. L., Vines, B. W., & Ramsay, J. (2007). Introduction to functional data analysis. *Canadian Psychology/Psychologie Canadienne*, 48(3), 135.
- Levy, B. J., & Anderson, M. C. (2002). Inhibitory processes and the control of memory retrieval. *Trends in Cognitive Sciences*, 6(7), 299–305.
- Lindauer, T. (1998). Attributive genitive constructions in German. In A. Alexiadou & C. Wilder (Eds.), *Possessors, predicates and movement in the determiner phrase* (pp. 109–140). Amsterdam: John Benjamins.
- Litvak, V., Mattout, J., Kiebel, S., Phillips, C., Henson, R., Kilner, J., ... Friston, K. (2011). EEG and MEG data analysis in SPM8. *Computational Intelligence and Neuroscience*, 2011, 852961.
- Liu, Z., & Guo, W. (2012). Functional mixed effects models. *Wiley Interdisciplinary Reviews: Computational Statistics*, 4(6), 527–534.
- Locker, L., Hoffman, L., & Bovaird, J. A. (2007). On the use of multilevel modeling as an alternative to items analysis in psycholinguistic research. *Behavior Research Methods*, 39(4), 723–730.
- Lockwood, H. T., & Macaulay, M. (2012). Prominence hierarchies. *Language and Linguistics Compass*, 6(7), 431–446.
- Luck, S. J. (2005a). *An introduction to the event-related potential technique*. Cambridge: MIT Press.
- Luck, S. J. (2005b). Ten simple rules for designing ERP experiments. In T. C. Handy (Ed.), *Event-related potentials: A methods handbook*. Cambridge: MIT Press.
- MacCallum, R. C., Zhang, S., Preacher, K. J., & Rucker, D. D. (2002). On the practice of dichotomization of quantitative variables. *Psychological Methods*, 7(1), 19–40.
- Makeig, S., Debener, S., Onton, J., & Delorme, A. (2004). Mining event-related brain dynamics. *Trends in Cognitive Sciences*, 8(5), 204–210.
- Malaia, E., Wilbur, R. B., & Weber-Fox, C. (2009). ERP evidence for telicity effects on syntactic processing in garden-path sentences. *Brain and Language*, 108(3), 145–

- Malaia, E., Wilbur, R. B., & Weber-Fox, C. (2011). Effects of verbal event structure on online thematic role assignment. *Journal of Psycholinguistic Research*, 41 (5), 323–345.
- Malaia, E., Wilbur, R. B., & Weber-Fox, C. (2013). Event end-point primes the Undergoer argument: Neurobiological bases of event structure processing. In B. Arsenijevic, B. Gehrke, & R. Marín. (Eds.), *Studies in the composition and decomposition of event predicates* (pp. 231–48). Springer.
- Malchukov, A. (2005). Case pattern splits, verb types and construction competition. In M. Amberber & H. de Hoop (Eds.), *Competition and variation in natural languages: The case for case* (pp. 73–118). Amsterdam: Elsevier.
- Manouilidou, C., & de Almeida, R. G. (2013). Processing correlates of verb typologies: Investigating internal structure and argument realization. *Linguistics*, 51 (4), 767–792.
- Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, 164 (1), 177–190.
- Martinez, J. G., Bohn, K. M., Carroll, R. J., & Morris, J. S. (2013). A study of Mexican free-tailed bat chirp syllables: Bayesian functional mixed modeling of nonstationary time series data with time-dependent spectra. *Journal of the American Statistical Association*, 108 (502), 514–526.
- Martín-Loeches, M., Ouyang, G., Rausch, P., Stürmer, B., Palazova, M., Schacht, A., & Sommer, W. (2017). Test–retest reliability of the N400 component in a sentence-reading paradigm. *Language, Cognition and Neuroscience*, 32 (10), 1261–1272.
- Matloff, N. (2011). *The art of R programming: A tour of statistical software design*. San Francisco: No Starch Press.
- McKoon, G., & Love, J. (2011). Verbs in the lexicon: Why is hitting easier than breaking? *Language and Cognition*, 3, 313–330.
- McKoon, G., & Macfarland, T. (2000). Externally and internally caused change of state verbs. *Language*, 76 (4), 833–858.
- McKoon, G., & Macfarland, T. (2002). Event templates in the lexical representations of verbs. *Cognitive Psychology*, 45 (1), 1–44.
- McKoon, G., & Ratcliff, R. (2003). Meaning through syntax: Language comprehension and the reduced relative clause construction. *Psychological Review*, 110 (3), 490–525.
- McRae, K., Ferretti, T. R., & Amyote, L. (1997). Thematic roles as verb-specific concepts. *Language and Cognitive Processes*, 12 (2/3), 137–176.
- Meyer, M. J., Coull, B. A., Versace, F., Cinciripini, P., & Morris, J. S. (2015). Bayesian function-on-function regression for multilevel functional data. *Biometrics*, 71 (3), 563–574. doi: 10.1111/biom.12299
- Morris, J. S. (2012). Statistical methods for proteomic biomarker discovery based on feature extraction or functional modeling approaches. *Statistics and Its Interface*, 5 (1), 117–136.
- Morris, J. S. (2015). Functional regression. *Annual Review of Statistics and Its Application*, 2, 321–359.
- Morris, J. S., Arroyo, C., Coull, B. A., Ryan, L. M., & Gortmaker, S. L. (2006). Wavelet-based

- functional mixed models to characterize population heterogeneity in accelerometer profiles: A case study. *Journal of the American Statistical Association*, 101 (476), 1352–1364.
- Morris, J. S., Baladandayuthapani, V., Herrick, R. C., Sanna, P., & Gutstein, H. (2011). Automated analysis of quantitative image data using isomorphic functional mixed models, with application to proteomics data. *The Annals of Applied Statistics*, 5(2A), 894–923.
- Morris, J. S., Brown, P. J., Baggerly, K. A., & Coombes, K. R. (2006). Analysis of mass spectrometry data using Bayesian wavelet-based functional mixed models. In K. A. Do, P. Mueller, & M. Vannucci (Eds.), *Bayesian inference for gene expression and proteomics* (p. 269–88). New York: Cambridge University Press.
- Morris, J. S., Brown, P. J., Herrick, R. C., Baggerly, K. A., & Coombes, K. R. (2008). Bayesian analysis of mass spectrometry proteomic data using wavelet-based functional mixed models. *Biometrics*, 64 (2), 479–489.
- Morris, J. S., & Carroll, R. J. (2006). Wavelet-based functional mixed models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 68 (2), 179–199.
- Morris, J. S., Coombes, K. R., Koomen, J., Baggerly, K. A., & Kobayashi, R. (2005). Feature extraction and quantification for mass spectrometry in biomedical applications using the mean spectrum. *Bioinformatics*, 21 (9), 1764–1775.
- Morris, J. S., & Herrick, R. C. (2012). WFMM user's guide [Computer software manual]. Retrieved from [https://biostatistics.mdanderson.org/SoftwareDownload/ProductSupportFiles/WFMM/wfmm\\_Users\\_Guide.pdf](https://biostatistics.mdanderson.org/SoftwareDownload/ProductSupportFiles/WFMM/wfmm_Users_Guide.pdf) (WFMM version 3.0)
- Næss, Å. (2004). What markedness marks: The markedness problem with direct objects. *Lingua*, 114 (9), 1186–1212.
- Næss, Å. (2007). *Prototypical transitivity*. Amsterdam: John Benjamins.
- Nason, G. (2013). wavethresh: Wavelets statistics and transforms. [Computer software manual]. Retrieved from <http://CRAN.R-project.org/package=wavethresh> (R package version 4.6.6)
- Neville, H., Nicol, J., Barss, A., Forster, K., & Garrett, M. (1991). Syntactically based sentence processing classes: Evidence from event-related brain potentials. *Journal of Cognitive Neuroscience*, 3 (2), 151–165.
- Nieuwenhuis, R., Grotenhuis, M. t., & Pelzer, B. (2012). influence.ME: Tools for detecting influential data in mixed effects models. *The R Journal*, 4 (2), 38–47.
- O'Bryan, E. L. (2003). *Event structure in language comprehension* (PhD Thesis). The University of Arizona.
- Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh Inventory. *Neuropsychologia*, 9 (1), 97–113.
- Olsen, S. (1991). Die deutsche Nominalphrase als "Determinansphrase". In S. Olsen & G. Fanselow (Eds.), *DET, COMP und INFL: Zur Syntax funktionaler Kategorien und grammatischer Funktionen* (pp. 35–56). Tübingen: Niemeyer.
- Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic

- anomaly. *Journal of Memory and Language*, 31 (6), 785–806.
- Paczynski, M., Jackendoff, R., & Kuperberg, G. R. (2014). When events change their nature: The neurocognitive mechanisms underlying aspectual coercion. *Journal of Cognitive Neuroscience*, 26 (9), 1905–1917.
- Palmer, M., Gildea, D., & Xue, N. (2010). Semantic role labeling. *Synthesis lectures on human language technologies*, Vol. 6. San Rafael: Morgan & Claypool Publishers.
- Perlmutter, D. M. (1978). Impersonal passives and the unaccusative hypothesis. In *Proceedings of the annual meeting of the Berkeley Linguistics Society* (Vol. 4).
- Pernet, C. R., Chauveau, N., Gaspar, C., & Rousselet, G. A. (2011). LIMO EEG: A toolbox for hierarchical Linear Modeling of ElectroEncephaloGraphic data. *Computational Intelligence and Neuroscience*, 2011, 3.
- Pernet, C. R., Sajda, P., & Rousselet, G. A. (2011). Single-trial analyses: Why bother? *Frontiers in Psychology*, 2, 322.
- Pickering, M. J., McElree, B., Frisson, S., Chen, L., & Traxler, M. J. (2006). Underspecification and aspectual coercion. *Discourse Processes*, 42 (2), 131–155.
- Piñango, M. M., Winnick, A., Ullah, R., & Zurif, E. (2006). Time-course of semantic composition: The case of aspectual coercion. *Journal of Psycholinguistic Research*, 35 (3), 233–244.
- Pinker, S. (1989). *Learnability and cognition: The acquisition of argument structure*. Cambridge: MIT Press.
- Pinker, S. (1994). How could a child use verb syntax to learn verb semantics? *Lingua*, 92, 377–410.
- Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology*, 118 (10), 2128–2148.
- Primus, B. (1999). *Cases and thematic roles: Ergative, accusative and active*. Tübingen: Niemeyer.
- Prince, A., & Smolensky, P. ([1993]/2004). *Optimality Theory: Constraint interaction in generative grammar*. Malden: Blackwell.
- Pylkkänen, L. (2008). Mismatching meanings in brain and behavior. *Language and Linguistics Compass*, 2 (4), 712–738.
- Pylkkänen, L., Brennan, J., & Bemis, D. K. (2011). Grounding the cognitive neuroscience of semantics in linguistic theory. *Language and Cognitive Processes*, 26 (9), 1317–1337.
- Pylkkänen, L., Martin, A. E., McElree, B., & Smart, A. (2009). The anterior midline field: Coercion or decision making? *Brain and Language*, 108 (3), 184–190.
- Pylkkänen, L., & McElree, B. (2007). An MEG study of silent meaning. *Journal of Cognitive Neuroscience*, 19 (11), 1905–1921.
- Pylkkänen, L., Oliveri, B., & Smart, A. J. (2009). Semantics vs. world knowledge in prefrontal cortex. *Language and Cognitive Processes*, 24 (9), 1313–1334.
- Quené, H., & van den Bergh, H. (2004). On multi-level modeling of data from repeated measures designs: A tutorial. *Speech Communication*, 43 (1), 103–121.
- R Development Core Team. (2013). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from <http://www.R-project>

- .org/
- Raaijmakers, J. G., Schrijnemakers, J., & Gremmen, F. (1999). How to deal with "the language-as-fixed-effect fallacy": Common misconceptions and alternative solutions. *Journal of Memory and Language*, 41(3), 416–426.
- Ramsay, J. O., & Silverman, B. W. (2005). *Functional data analysis* (Second ed.). New York: Springer.
- Rappaport Hovav, M., & Levin, B. (1998). Building verb meanings. In M. Butt & W. Geuder (Eds.), *The projection of arguments: Lexical and compositional factors* (pp. 97–134). Stanford: CSLI Publications.
- Rappaport Hovav, M., & Levin, B. (2001). An event structure account of English resultatives. *Language*, 77(4), 766–797.
- Rappaport Hovav, M., & Levin, B. (2005). Change of state verbs: implications for theories of argument projection. In N. Erteschik-Shir & T. R. Rapoport (Eds.), *The syntax of aspect. Deriving thematic and aspectual interpretation* (pp. 274–286). Oxford: Oxford University Press.
- Rayner, K., & Duffy, S. A. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory & Cognition*, 14(3), 191–201.
- Raz, J., Dickerson, L., & Turetsky, B. (1999). A wavelet packet model of evoked potentials. *Brain and Language*, 66(1), 61–88.
- Reinhart, T. (2003). The theta system – an overview. *Theoretical Linguistics*, 28(3), 229–290.
- Roehm, D., Bornkessel-Schlesewsky, I., & Schlewsky, M. (2007). The internal structure of the N400: Frequency characteristics of a language-related ERP component. *Chaos and Complexity Letters*, 2(2/3), 365–395.
- Roehm, D., Schlewsky, M., Bornkessel, I., Frisch, S., & Haider, H. (2004). Fractionating language comprehension via frequency characteristics of the human EEG. *Neuroreport*, 15(3), 409–412.
- Romagno, D., Rota, G., Ricciardi, E., & Pietrini, P. (2012). Where the brain appreciates the final state of an event: The neural correlates of telicity. *Brain and Language*, 123(1), 68–74.
- Rosch, E. H. (1973). On the internal structure of perceptual and semantic categories. In T. E. Moore (Ed.), *Cognitive development and the acquisition of language* (pp. 111–144). New York: Academic Press.
- Rossion, B., & Jacques, C. (2008). Does physical interstimulus variance account for early electrophysiological face sensitive responses in the human brain? Ten lessons on the N170. *NeuroImage*, 39(4), 1959–1979.
- Rousselet, G. A., & Pernet, C. R. (2011). Quantifying the time course of visual object processing using ERPs: It's time to up the game. *Frontiers in Psychology*, 2, 107.
- Rozwadowska, B. (1988). Thematic restrictions on derived nominals in thematic relations. In W. Wilkins (Ed.), *Syntax and semantics 21: Thematic relations* (pp. 147–165). San Diego: Academic Press.



- Ruppert, D., Wand, M. P., & Carroll, R. J. (2003). *Semiparametric regression*. Cambridge: Cambridge University Press.
- Samar, V. J., Bopardikar, A., Rao, R., & Swartz, K. (1999). Wavelet analysis of neuroelectric waveforms: A conceptual tutorial. *Brain and Language*, 66(1), 7–60.
- Schacht, A., Sommer, W., Shmulevich, O., Martíenz, P. C., & Martín-Loeches, M. (2014). Differential task effects on N400 and P600 elicited by semantic and syntactic violations. *PLOS ONE*, 9(3), e91226.
- Schielzeth, H., & Forstmeier, W. (2009). Conclusions beyond support: Overconfident estimates in mixed models. *Behavioral Ecology*, 20(2), 416–420.
- Siegle, G. J., Ichikawa, N., & Steinhauer, S. (2008). Blink before and after you think: Blinks occur prior to and following cognitive load indexed by pupillary responses. *Psychophysiology*, 45(5), 679–687.
- Smith, N. J., & Kutas, M. (2014a). Regression-based estimation of ERP waveforms: II. Non-linear effects, overlap correction, and practical considerations. *Psychophysiology*.
- Smith, N. J., & Kutas, M. (2014b). Regression-based estimation of ERP waveforms: I. The rERP framework. *Psychophysiology*.
- Solomyak, O., & Marantz, A. (2009). Lexical access in early stages of visual word processing: A single-trial correlational MEG study of heteronym recognition. *Brain and Language*, 108(3), 191–196.
- Solstad, T. (2010). Post-nominal genitives and prepositional phrases in German: A uniform analysis. In A. Alexiadou & M. Rathert (Eds.), *The syntax of nominalizations across languages and frameworks* (pp. 219–251). Berlin: de Gruyter.
- Sorace, A. (2000). Gradients in auxiliary selection with intransitive verbs. *Language*, 76(4), 859–890.
- Sorace, A., & Keller, F. (2005). Gradience in linguistic data. *Lingua*, 115(11), 1497–1524.
- Steen, J. (2010). *An analysis of ERP data by wavelet-based functional mixed effect modelling* (Master's thesis). Ghent University.
- Steinhauer, K., Pancheva, R., Newman, A. J., Gennari, S., & Ullman, M. T. (2001). How the mass counts: An electrophysiological approach to the processing of lexical features. *Neuroreport*, 12(5), 999–1005.
- Stiebels, B. (2006). From rags to riches. Nominal linking in contrast to verbal linking. In D. Wunderlich (Ed.), *Advances in the theory of the lexicon* (pp. 167–234). Berlin: de Gruyter.
- Stockall, L., Husband, E. M., & Beretta, A. (2010). The online composition of events. *Queen Mary's Occasional Papers Advancing Linguistics*, 19.
- Stowell, T. (1981). *Origins of phrase structure* (PhD thesis). MIT.
- Talmy, L. (1985). Lexicalization patterns: Semantic structure in lexical forms. In T. Shopen (Ed.), *Language typology and syntactic description 3: Grammatical categories and the lexicon* (pp. 57–149). Cambridge: Cambridge University Press.
- Tanenhaus, M. K., Carlson, G., & Trueswell, J. C. (1989). The role of thematic structures in interpretation and parsing. *Language and Cognitive Processes*, 4(3-4), S1211–S1234.
- Tenny, C. (1992). The aspectual interface hypothesis. In I. A. Sag & A. Szabolcsi (Eds.),

- Lexical matters* (pp. 1–27). Stanford: CSLI Publications.
- Tenny, C. (1994). *Aspectual roles and the syntax-semantics interface*. Dordrecht: Kluwer.
- Testelec, Y. G. (1998). On two parameters of transitivity. In L. Kulikov & H. Vater (Eds.), *Typology of verbal categories: Papers presented to Vladimir Nedjalkov on the occasion of his 70th birthday* (pp. 29–45). Tübingen: Niemeyer.
- The MathWorks Inc. (2012). *MATLAB® version 7.14 (R2012a)*. Natick, Massachusetts.
- Thierry, G., Martin, C. D., Downing, P., & Pegna, A. J. (2007). Controlling for interstimulus perceptual variance abolishes N170 face selectivity. *Nature Neuroscience*, 10(4), 505–511.
- Thivierge, J.-P. (2008). Higher derivatives of ERP responses to cross-modality processing. *Neuroinformatics*, 6(1), 35–46.
- Thompson-Schill, S. L., Bedny, M., & Goldberg, R. F. (2005). The frontal lobes and the regulation of mental activity. *Current Opinion in Neurobiology*, 15(2), 219–224.
- Thompson-Schill, S. L., & Botvinick, M. M. (2006). Resolving conflict: A response to Martin and Cheng (2006). *Psychonomic Bulletin & Review*, 13(3), 402–408.
- Thornhill, D. E., & Van Petten, C. (2012). Lexical versus conceptual anticipation during sentence processing: Frontal positivity and N400 ERP components. *International Journal of Psychophysiology*, 83, 382–392.
- Tremblay, A., & Baayen, R. H. (2010). Holistic processing of regular four-word sequences: A behavioral and ERP study of the effects of structure, frequency, and probability on immediate free recall. In D. Wood (Ed.), *Perspectives on formulaic language: Acquisition and communication* (pp. 151–73). London: The Continuum International Publishing Group.
- Tremblay, A., & Ransijn, J. (2013). LMERConvenienceFunctions: A suite of functions to back-fit fixed effects and forward-fit random effects, as well as other miscellaneous functions. [Computer software manual]. Retrieved from <http://CRAN.R-project.org/package=LMERConvenienceFunctions> (R package version 2.5)
- Trueswell, J. C., Tanenhaus, M. K., & Garnsey, S. M. (1994). Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal of Memory and Language*, 33(3), 285–318.
- Tsunoda, T. (1981). Split case-marking patterns in verb-types and tense/aspect/mood. *Linguistics*, 19, 389–438.
- Tsunoda, T. (1985). Remarks on transitivity. *Journal of Linguistics*, 21(2), 385–396.
- Van Berkum, J. J. A. (2012). The electrophysiology of discourse and conversation. In M. Spivey, K. McRae, & M. Joannisse (Eds.), *The Cambridge handbook of psycholinguistics* (pp. 589–614). Cambridge: Cambridge University Press.
- Van Berkum, J. J. A., Brown, C. M., & Hagoort, P. (1999). Early referential context effects in sentence processing: Evidence from event-related brain potentials. *Journal of Memory and Language*, 41(2), 147–182.
- Van Berkum, J. J. A., Koornneef, A. W., Otten, M., & Nieuwland, M. S. (2007). Establishing reference in language comprehension: An electrophysiological perspective. *Brain Research*, 1146, 158–171.

- Van Dyke, J. A., & McElree, B. (2006). Retrieval interference in sentence comprehension. *Journal of Memory and Language*, 55(2), 157–166.
- Van Herten, M., Kolk, H. H. J., & Chwilla, D. J. (2005). An ERP study of P600 effects elicited by semantic anomalies. *Cognitive Brain Research*, 22(2), 241–255.
- Van Petten, C., & Luka, B. J. (2012). Prediction during language comprehension: Benefits, costs, and ERP components. *International Journal of Psychophysiology*, 83, 176–190.
- Van Valin Jr., R. D. (1990). Semantic parameters of split intransitivity. *Language*, 66(2), 221–260.
- Van Valin Jr., R. D., & LaPolla, R. J. (1997). *Syntax: Structure, meaning, and function*. Cambridge: Cambridge University Press.
- Vasishth, S. (2006). On the proper treatment of spillover in real-time reading studies: Consequences for psycholinguistic theories. In *Proceedings of the International Conference on Linguistic Evidence* (pp. 96–100). Tübingen, Germany.
- Vendler, Z. (1957). Verbs and times. *The Philosophical Review*, 56, 143–160.
- Vigliocco, G., Vinson, D. P., Druks, J., Barber, H., & Cappa, S. F. (2011). Nouns and verbs in the brain: A review of behavioural, electrophysiological, neuropsychological and imaging studies. *Neuroscience & Biobehavioral Reviews*, 35(3), 407–426.
- Vines, B. W., Nuzzo, R. L., & Levitin, D. J. (2005). Analyzing temporal dynamics in music: Differential calculus, physics, and functional data analysis techniques. *Music Perception: An Interdisciplinary Journal*, 23(2), 137–152.
- Wang, L., Bastiaansen, M., Yang, Y., & Hagoort, P. (2012). Information structure influences depth of syntactic processing: Event-related potential evidence for the Chomsky illusion. *PloS ONE*, 7(10), e47917.
- Weckerly, J., & Kutas, M. (1999). An electrophysiological analysis of animacy effects in the processing of object relative sentences. *Psychophysiology*, 36(5), 559–570.
- Wickham, H. (2009). *ggplot2: Elegant graphics for data analysis*. New York: Springer.
- Wickham, H., & Francois, R. (2014). dplyr: a grammar of data manipulation [Computer software manual]. Retrieved from <http://CRAN.R-project.org/package=dplyr> (R package version 0.1.3)
- Wike, E. L., & Church, J. D. (1976). Comments on Clark's "The language-as-fixed-effect fallacy". *Journal of Verbal Learning and Verbal Behavior*, 15(3), 249–255.
- Wood, S. (2006). *Generalized additive models: An introduction with R*. Chapman & Hall/CRC Press.
- Woolrich, M. W., Beckmann, C. F., Nichols, T. E., & Smith, S. M. (2009). Statistical analysis of fMRI data. In M. Filippi (Ed.), *fMRI techniques and protocols* (pp. 179–236). New York: Springer.
- Worsley, K. (2003). Detecting activation in fMRI data. *Statistical Methods in Medical Research*, 12(5), 401–418.
- Wunderlich, D. (1996). Lexical categories. *Theoretical Linguistics*, 22(1-2), 1–48.
- Wunderlich, D. (1997). Cause and the structure of verbs. *Linguistic Inquiry*, 28(1), 27–68.
- Xie, Y. (2013). animation: An R package for creating animations and demonstrating statistical

- methods. *Journal of Statistical Software*, 53(1), 1–27.
- Zhang, L., Baladandayuthapani, V., Zhu, H., Baggerly, K. A., Majewski, T., Czerniak, B., & Morris, J. S. (2015). Functional CAR models for spatially correlated functional datasets. *Journal of the American Statistical Association*.
- Zhu, H., Brown, P. J., & Morris, J. S. (2011). Robust, adaptive functional regression in functional mixed model framework. *Journal of the American Statistical Association*, 106(495), 1167–1179.
- Zhu, H., Brown, P. J., & Morris, J. S. (2012). Robust classification of functional and quantitative image data using functional mixed models. *Biometrics*, 68, 1260–1268.
- Zhu, H., Versace, F., Cinciripini, P. M., Rausch, P., & Morris, J. S. (accepted). Robust and Gaussian spatial functional regression models for analysis of event-related potentials. *NeuroImage*.

## Appendix A

### Nominalisation infos

The following table lists all nominalisation used for the construction of experimental sentences in the three experiments conducted together with the most respective AI value (as rated in the pretests – see sections 2.7.2 and 4.2.1), the base verb in German and English translations thereof. Since use of nominalisations varied slightly across experiments, the table also shows which nominalisation was used in which experiment(s).

**Table A.1:** Infos on nominalisations used in the three experiments. Abbreviations: AI = Affectedness Index; Accept. rating = Acceptability rating; SPR = Self-paced reading; ERP = Event-related potentials.

Nominalisation	AI	Used in experiment			Base verb	
		Accept. rating	SPR	ERP	German	English translation(s)
Ablehnung	4.72	✓	✓	✓	ablehnen	reject so.
Abschirmung	3.72	✓	✓	✓	abschirmen	screen off so.
Absetzung	6	✓	✓	✓	absetzen	dismiss so.
Abwerbung	4.76	✓	✓	✓	abwerben	entice so. away
Anfeindung	3.68	✓	✓	✗	anfeinden	be hostile to so.
Anhörung	2.92	✓	✓	✓	anhören	hear so. (a party)
Aufforderung	2.09	✗	✗	✓	auffordern	request so. (to do sthg.)
Aufmunterung	4.84	✓	✓	✓	aufmuntern	cheer up so.
Ausbeutung	4.84	✓	✓	✓	ausbeuten	exploit so.
Ausbildung	5.68	✓	✓	✓	ausbilden	educate, train so.
Auslieferung	5.68	✓	✓	✓	ausliefern	extradite so.
Auswechslung	5.88	✓	✓	✓	auswechseln	replace, substitute so.
Ausweisung	6.28	✓	✓	✓	ausweisen	expel so.
Beförderung	5.56	✓	✓	✓	befördern	promote so.
Befragung	2.68	✓	✓	✓	befragen	interview, question so.
Befreiung	6.28	✓	✓	✓	befreien	free so.
Begrüßung	2.8	✓	✓	✓	begrüßen	greet so.
Behandlung	4.68	✓	✓	✓	behandeln	treat so.

Bekehrung	5.16	✓	✓	✓	bekehren	convert so.
Belagerung	3.72	✓	✓	✓	belagern	besiege so.
Belästigung	4.44	✓	✓	✓	belästigen	harass so.
Belehrung	4.2	✓	✓	✓	belehren	instruct, mirandise so.
Beleidigung	3.44	✓	✓	✓	beleidigen	insult so.
Beratung	2.8	✓	✓	✓	beraten	advise so.
Bergung	5.56	✓	✓	✓	bergen	retrieve, recover so.
Berührung	1.91	X	X	✓	berühren	touch so.
Beschattung	2.56	✓	✓	✓	beschatten	tail so.
Beschimpfung	3.8	✓	✓	✓	beschimpfen	abuse so.
Beseitigung	6.6	✓	✓	✓	beseitigen	remove so.
Bespitzelung	3.36	✓	✓	X	bespitzeln	spy on so.
Bestechung	4.88	✓	✓	✓	bestechen	bribe so.
Bestrafung	5.24	✓	✓	✓	bestrafen	punish, penalise so.
Betreuung	3.36	✓	✓	✓	betreuen	take care of so., attend to so.
Bevormundung	4.12	✓	✓	✓	bevormunden	domineer, patronise so.
Bewachung	3.24	✓	✓	✓	bewachen	guard so.
Bewunderung	1.8	✓	✓	✓	bewundern	admire so.
Diskriminierung	4.68	✓	✓	✓	diskriminieren	discriminate against so.
Durchleuchtung	3.16	✓	✓	✓	durchleuchten	scrutinise, check so. thoroughly
Durchsuchung	3.56	✓	✓	✓	durchsuchen	frisk so.
Einbürgerung	5.32	✓	✓	X	einbürgern	naturalise so.
Einlieferung	5.48	✓	✓	✓	einliefern	hospitalise so.
Einschätzung	1.91	X	X	✓	einschätzen	assess so.
Einschüchterung	4.52	✓	✓	✓	einschüchtern	intimidate so.
Enteignung	6.16	✓	✓	✓	enteignen	dispossess so.
Enterbung	5.48	✓	✓	✓	enterben	disinherit so.
Entführung	6.08	✓	✓	✓	entführen	abduct, kidnap so.
Entgiftung	6.12	✓	X	X	entgiften	to detoxify so.
Enthauptung	7	✓	✓	✓	enthaupten	decapitate so.
Enthebung	6	✓	✓	X	entheben	displace, impeach so.
Entlarvung	5.68	✓	✓	✓	entlarven	expose so.
Entlassung	6.12	✓	✓	✓	entlassen	dismiss so.
Entlastung	4.72	✓	✓	✓	entlasten	relieve, disburden so.
Entmachtung	6.08	✓	✓	✓	entmachten	disempower so.
Entwaffnung	5.8	✓	✓	✓	entwaffen	disarm so.
Ergreifung	6.08	✓	✓	✓	ergreifen	apprehend, catch so.
Ermahnung	3.56	✓	✓	✓	ermahnen	admonish so.
Ermordung	6.76	✓	✓	✓	ermorden	murder so.
Ermutigung	4.52	✓	✓	✓	ermutigen	encourage so.
Erpressung	4.92	✓	✓	✓	erpressen	blackmail so.
Förderung	5	✓	✓	✓	fördern	support, patronise so.

Freilassung	6	✓	✓	✓	freilassen	release so.
Heilung	6.44	✓	✓	✓	heilen	cure so.
Inhaftierung	6	✓	✓	✓	inhaftieren	detain, arrest so.
Manipulierung	4.48	✓	✓	✓	manipulieren	manipulate so.
Rehabilitierung	5.48	✓	✓	✓	rehabilitieren	rehabilitate so.
Rettung	6.24	✓	✓	✓	retten	rescue so.
Schulung	5.16	✓	✓	✓	schulen	instruct, train so.
Täuschung	3.32	✓	✓	✓	täuschen	deceive so.
Tötung	6.76	✓	✓	✓	töten	kill so.
Überprüfung	3.24	✓	✓	✓	überprüfen	check up on so.
Überstellung	4.68	✓	✓	✓	überstellen	transfer so.
Überwachung	2.88	✓	✓	✓	überwachen	watch so., keep so. under surveillance
Umarmung	2.84	✓	✓	✓	umarmen	hug so.
Unterdrückung	4.92	✓	✓	✓	unterdrücken	oppress, subdue so.
Unterstützung	3.72	✓	✓	✓	unterstützen	support so.
Untersuchung	3.52	✓	✓	✓	untersuchen	examine so.
Verachtung	1.64	X	X	✓	verachten	despise so.
Verbannung	6	✓	✓	✓	verbannen	banish, expulse so.
Verehrung	1.73	X	X	✓	verehren	admire so.
Verhaftung	6.28	✓	✓	✓	verhaften	arrest, detain so.
Verlegung	5.48	✓	✓	✓	verlegen	transfer, relocate so.
Vernehmung	3.68	✓	✓	✓	vernehmen	question so.
Verschleppung	6.16	✓	✓	✓	verschleppen	haul so. off
Versetzung	5.36	✓	✓	✓	versetzen	relocate, transfer so.
Versorgung	4.76	✓	✓	✓	versorgen	tend to so., provide for so.
Verteidigung	3.4	✓	✓	✓	verteidigen	defend so.
Vertreibung	6.12	✓	✓	✓	vertreiben	expel, dislodge so.
Verurteilung	6.04	✓	✓	✓	verurteilen	convict so.
Würdigung	3.8	✓	✓	✓	würdigen	dignify, honour so.
Zurückweisung	4.92	✓	✓	X	zurückweisen	reject so.





## Appendix B

# Supplementary materials for the acceptability rating study

### Stimulus sentences

The following pages list the stimulus sentences used in the experiment, alphabetically ordered by the relevant *-ung*-nominalisation and numbered consecutively by 'story'. The first sentence represents the context sentence, the second the continuation sentence; GenO and GenS subscripts indicate the classifications of the genitives presented in the continuation sentences.

#### Ablehnung

1. *Context*: Der Direktor weigerte sich, den Forscher einzustellen.  
*Continuation*: Die Ablehnung des Direktors<sub>GenS</sub>/Forschers<sub>GenO</sub> enttäuschte die Kollegen im Labor.
2. *Context*: Der Professor verweigerte die Aufnahme des talentierten Bewerbers an die Schauspielschule.  
*Continuation*: Die Ablehnung des Professors<sub>GenS</sub>/Bewerbers<sub>GenO</sub> überraschte die Kollegen an der Uni.

#### Abschirmung

3. *Context*: Der Bodyguard ließ während des gesamten Termins keine Fans an den Sportler heran.  
*Continuation*: Die Abschirmung des Bodyguards<sub>GenS</sub>/Sportlers<sub>GenO</sub> frustrierte die Fans nach wenigen Minuten.
4. *Context*: Der Anwalt hielt sämtliche Reporter von dem Opfer fern.  
*Continuation*: Die Abschirmung des Anwalts<sub>GenS</sub>/Opfers<sub>GenO</sub> ermöglichte die Vorbereitung auf den Prozess.

#### Absetzung

5. *Context*: Der Kanzler entließ seinen Minister, nachdem der Skandal aufgefliegen war.  
*Continuation*: Die Absetzung des Kanzlers<sub>GenS</sub>/Ministers<sub>GenO</sub> rettete die Regierung vor einer Krise.
6. *Context*: Der Admiral suspendierte den Kapitän wegen der schweren Vorwürfe.  
*Continuation*: Die Absetzung des Admirals<sub>GenS</sub>/Kapitäns<sub>GenO</sub> brachte den Minister aus der Schusslinie der Medien.

#### Abwerbung

7. *Context:* Der Manager des Fußballclubs überzeugte den Star der Konkurrenz mit einem millionenschweren Vertrag, zu wechseln.  
*Continuation:* Die Abwerbung des Managers<sub>GenS</sub>/Stars<sub>GenO</sub> verärgerte die Fans nach der erfolglosen Saison.
8. *Context:* Der Chef des Konzerns brachte den Forscher dazu, in sein Unternehmen zu wechseln.  
*Continuation:* Die Abwerbung des Chefs<sub>GenS</sub>/Forschers<sub>GenO</sub> hatte ein Nachspiel vor Gericht.

#### Anfeindung

9. *Context:* Der Politiker beschimpfte den Reporter während der Aufzeichnung des Interviews.  
*Continuation:* Die Anfeindung des Politikers<sub>GenS</sub>/Reporters<sub>GenO</sub> schockierte das Publikum im Studio.
10. *Context:* Der Professor attackierte den Moderator während einer Diskussion verbal.  
*Continuation:* Die Anfeindung des Professors<sub>GenS</sub>/Moderators<sub>GenO</sub> irritierte die Zuhörer im Saal.

#### Anhörung

11. *Context:* Der Senat befragte den Minister zu den Vorgängen.  
*Continuation:* Die Anhörung des Senats<sub>GenS</sub>/Ministers<sub>GenO</sub> beschäftigte die Journalisten für einige Tage.
12. *Context:* Der Vorstand befragte den Manager, um seine Sicht der Dinge zu hören.  
*Continuation:* Die Anhörung des Vorstands<sub>GenS</sub>/Managers<sub>GenO</sub> erbrachte keine Hinweise auf einen Betrug.

#### Aufmunterung

13. *Context:* Der Vater tröstete das Kind nach der verpatzten Klausur.  
*Continuation:* Die Aufmunterung des Vaters<sub>GenS</sub>/Kindes<sub>GenO</sub> rettete das Wochenende für die ganze Familie.
14. *Context:* Der Clown brachte den Patienten nach langer Zeit wieder zum Lachen.  
*Continuation:* Die Aufmunterung des Clowns<sub>GenS</sub>/Patienten<sub>GenO</sub> freute die Angehörigen nach den schweren Wochen.

#### Ausbeutung

15. *Context:* Der Professor übertrug dem Assistenten jahrelang seine eigene Arbeit.  
*Continuation:* Die Ausbeutung des Professors<sub>GenS</sub>/Assistenten<sub>GenO</sub> verärgerte viele Kollegen am Institut.
16. *Context:* Der Wirt ließ seinen Kellner für zu wenig Geld zu viel arbeiten.  
*Continuation:* Die Ausbeutung des Wirts<sub>GenS</sub>/Kellners<sub>GenO</sub> brachte den Vater auf die Palme.

#### Ausbildung

17. *Context:* Der Meister vermittelte dem Lehrling sein ganzes Können mit viel Begeisterung.  
*Continuation:* Die Ausbildung des Meisters<sub>GenS</sub>/Lehrlings<sub>GenO</sub> beeindruckte die Eltern wegen ihrer Gründlichkeit.
18. *Context:* Der Koch brachte dem Azubi im Lauf der Zeit viele neue Sachen bei.  
*Continuation:* Die Ausbildung des Kochs<sub>GenS</sub>/Azubis<sub>GenO</sub> beeindruckte den Chef von dem kleinen Restaurant.

#### Auslieferung

19. *Context:* Der Senat schickte den geflohenen Diktator zurück in sein Land.  
*Continuation:* Die Auslieferung des Senats<sub>GenS</sub>/Diktators<sub>GenO</sub> verbesserte die Stimmung zwischen den Staaten.
20. *Context:* Der Richter überstellte den Verdächtigen nach langem Zögern in seine Heimat.  
*Continuation:* Die Auslieferung des Richters<sub>GenS</sub>/Verdächtigen<sub>GenO</sub> verärgerte den Staatsanwalt am zuständigen Gericht.

#### Auswechslung

21. *Context:* Der Trainer nahm den Stürmer kurz nach der Halbzeitpause vom Feld.  
*Continuation:* Die Auswechslung des Trainers<sub>GenS</sub>/Stürmers<sub>GenO</sub> veranlasste die Fans zu einem Pfeifkonzert.

22. *Context:* Der Kläger suchte sich kurz vor dem Prozess einen neuen Anwalt.  
*Continuation:* Die Auswechslung des Klägers<sub>GenS</sub>/Anwalts<sub>GenO</sub> verzögerte den Prozessbeginn um einige Wochen.

#### **Ausweisung**

23. *Context:* Der Beamte lehnte den Antrag des Flüchtlings auf Asyl ab.  
*Continuation:* Die Ausweisung des Beamten<sub>GenS</sub>/Flüchtlings<sub>GenO</sub> schockierte viele Leute wegen ihrer Härte.
24. *Context:* Der Richter entschied, den Asylanten in sein Heimatland zurückzuschicken.  
*Continuation:* Die Ausweisung des Richters<sub>GenS</sub>/Asylanten<sub>GenO</sub> erschütterte die Zuschauer im Saal.

#### **Beförderung**

25. *Context:* Der Manager ernannte den Assistenten zum Projektleiter.  
*Continuation:* Die Beförderung des Managers<sub>GenS</sub>/Assistenten<sub>GenO</sub> freute die Kollegen im Büro.
26. *Context:* Der Minister machte den Kapitän wegen seiner Verdienste zum Admiral.  
*Continuation:* Die Beförderung des Ministers<sub>GenS</sub>/Kapitäns<sub>GenO</sub> freute die Kollegen am Stützpunkt.

#### **Befragung**

27. *Context:* Der Polizist verhörte den Zeugen zum Ablauf des Überfalls.  
*Continuation:* Die Befragung des Polizisten<sub>GenS</sub>/Zeugen<sub>GenO</sub> brachte einen Hinweis auf das Fluchtauto.
28. *Context:* Der Ermittler erhielt von dem Opfer beim dritten Treffen wichtige Hinweise auf den Täter.  
*Continuation:* Die Befragung des Ermittlers<sub>GenS</sub>/Opfers<sub>GenO</sub> brachte die Wende in dem Fall.

#### **Befreiung**

29. *Context:* Der Polizist rettete den Millionär aus der Geiselhaft.  
*Continuation:* Die Befreiung des Polizisten<sub>GenS</sub>/Millionärs<sub>GenO</sub> erlöste die Familie nach Wochen der Ungewissheit.
30. *Context:* Der Kommissar holte den verschleppten Jungen aus dem Keller des Entführers.  
*Continuation:* Die Befreiung des Kommissars<sub>GenS</sub>/Jungen<sub>GenO</sub> brachte den Fall in die internationalen Medien.

#### **Begrüßung**

31. *Context:* Der Kanzler empfing den Papst bei der Ankunft am Flughafen.  
*Continuation:* Die Begrüßung des Kanzlers<sub>GenS</sub>/Papstes<sub>GenO</sub> folgte dem Protokoll auf strikte Weise.
32. *Context:* Der Rektor hieß den Professor in einer Rede willkommen.  
*Continuation:* Die Begrüßung des Rektors<sub>GenS</sub>/Professors<sub>GenO</sub> amüsierte die Gäste im Festsaal der Uni.

#### **Behandlung**

33. *Context:* Der Arzt therapierte den Patienten mit einer neuen Methode.  
*Continuation:* Die Behandlung des Arztes<sub>GenS</sub>/Patienten<sub>GenO</sub> weckte neue Hoffnung bei den Angehörigen.
34. *Context:* Der Therapeut massierte den Sportler mit einer neuen Technik.  
*Continuation:* Die Behandlung des Therapeuten<sub>GenS</sub>/Sportlers<sub>GenO</sub> verringerte die Schmerzen nach wenigen Sitzungen.

#### **Bekehrung**

35. *Context:* Der Priester überzeugte den Häftling, in die Kirche einzutreten.  
*Continuation:* Die Bekehrung des Priesters<sub>GenS</sub>/Häftlings<sub>GenO</sub> brachte den Direktor zum Staunen.
36. *Context:* Laut der Legende brachte der Bischof den Sünder zurück auf den Weg des Glaubens.  
*Continuation:* Die Bekehrung des Bischofs<sub>GenS</sub>/Sünders<sub>GenO</sub> inspirierte den Dichter zu seinem Lied.

#### **Belagerung**

37. *Context*: Der Fan campierte monatelang vor dem Grundstück des Sängers.  
*Continuation*: Die Belagerung des Fans<sub>GenS</sub>/Sängers<sub>GenO</sub> zwang die Polizei nach einiger Zeit zum Handeln.
38. *Context*: Das Heer des Königs umzingelte das Lager des Rebellen für mehrere Wochen.  
*Continuation*: Die Belagerung des Königs<sub>GenS</sub>/Rebellen<sub>GenO</sub> entschied den Krieg nach zwei Jahren.

#### **Belästigung**

39. *Context*: Ein Betrunkener provozierte den Wächter in der U-Bahn-Station.  
*Continuation*: Die Belästigung des Betrunkenen<sub>GenS</sub>/Wächters<sub>GenO</sub> alarmierte einen Fahrgast auf dem Bahnsteig.
40. *Context*: Der Junge bedrängte das Mädchen mitten im Park.  
*Continuation*: Die Belästigung des Jungen<sub>GenS</sub>/Mädchens<sub>GenO</sub> veranlasste einen Jogger zum Eingreifen.

#### **Belehrung**

41. *Context*: Der Lehrer ermahnte den Schüler wegen seines schlechten Benehmens.  
*Continuation*: Die Belehrung des Lehrers<sub>GenS</sub>/Schülers<sub>GenO</sub> amüsierte die Mitschüler für einige Minuten.
42. *Context*: Der Anwalt klärte den Angeklagten vor dem Prozess über seine Rechte auf.  
*Continuation*: Die Belehrung des Anwalts<sub>GenS</sub>/Angeklagten<sub>GenO</sub> legte die Basis für die weitere Zusammenarbeit.

#### **Beleidigung**

43. *Context*: Der Schüler beschimpfte den Lehrer in der Pause.  
*Continuation*: Die Beleidigung des Schülers<sub>GenS</sub>/Lehrers<sub>GenO</sub> schockierte die Mitschüler in der Klasse.
44. *Context*: Der Fahrer zeigte dem Polizisten während einer Kontrolle den Mittelfinger.  
*Continuation*: Die Beleidigung des Fahrers<sub>GenS</sub>/Polizisten<sub>GenO</sub> hatte ein Nachspiel vor Gericht.

#### **Beratung**

45. *Context*: Der Anwalt half dem Millionär, eine Menge Steuern zu sparen.  
*Continuation*: Die Beratung des Anwalts<sub>GenS</sub>/Millionärs<sub>GenO</sub> brachte weitere Kunden aus wohlhabenden Kreisen.
46. *Context*: Der externe Experte der PR-Agentur half dem Kandidaten während des Wahlkampfs mit seiner Strategie.  
*Continuation*: Die Beratung des Experten<sub>GenS</sub>/Kandidaten<sub>GenO</sub> brachte viel Geld in die Kasse der Agentur.

#### **Bergung**

47. *Context*: Der Soldat zog den Touristen aus einer Gletscherspalte.  
*Continuation*: Die Bergung des Soldaten<sub>GenS</sub>/Touristen<sub>GenO</sub> beschäftigte die Reporter aus der Umgebung.
48. *Context*: Der Kapitän holte den verunglückten Fischer aus dem kalten Wasser.  
*Continuation*: Die Bergung des Kapitäns<sub>GenS</sub>/Fischers<sub>GenO</sub> beendete die Suche nach einigen Stunden.

#### **Beschattung**

49. *Context*: Der Agent beobachtete den Politiker über mehrere Wochen.  
*Continuation*: Die Beschattung des Agenten<sub>GenS</sub>/Politikers<sub>GenO</sub> brachte viele Informationen für die geplante Aktion.
50. *Context*: Der Geheimdienst ließ den Drogendealer während seines Aufenthalts nicht aus den Augen.  
*Continuation*: Die Beschattung des Geheimdienstes<sub>GenS</sub>/Drogendealers<sub>GenO</sub> erbrachte genügend Beweise für eine Verhaftung.

#### **Beschimpfung**

51. *Context*: Der Fahrer beleidigte den Polizisten bei einer Verkehrskontrolle.  
*Continuation*: Die Beschimpfung des Fahrers<sub>GenS</sub>/Polizisten<sub>GenO</sub> veranlasste einen Kollegen zu einer Anzeige.

52. *Context*: Der Chef schockierte den Arbeiter in der Besprechung mit seinen verletzenden Bemerkungen.  
*Continuation*: Die Beschimpfung des Chefs<sub>GenS</sub>/Arbeiters<sub>GenO</sub> verschlechterte die Stimmung während der Besprechung noch weiter.

#### **Beseitigung**

53. *Context*: Der Agent vergiftete den Diktator beim Abendessen.  
*Continuation*: Die Beseitigung des Agenten<sub>GenS</sub>/Diktators<sub>GenO</sub> öffnete den Weg zu einem Machtwechsel.
54. *Context*: Der Manager der Firma trieb den Konkurrenten gezielt in den Konkurs.  
*Continuation*: Die Beseitigung des Managers<sub>GenS</sub>/Konkurrenten<sub>GenO</sub> machte die Firma zum Marktführer.

#### **Bespitzelung**

55. *Context*: Der Spion hörte das Telefon des Ministers ab.  
*Continuation*: Die Bespitzelung des Spions<sub>GenS</sub>/Ministers<sub>GenO</sub> lieferte viele Informationen für die bevorstehende Verhandlung.
56. *Context*: Der Vorstand überprüfte die Telefonate des Managers.  
*Continuation*: Die Bespitzelung des Vorstands<sub>GenS</sub>/Managers<sub>GenO</sub> verursachte einen Skandal im Unternehmen.

#### **Bestechung**

57. *Context*: Der Unternehmer versprach dem Beamten eine Menge Geld, wenn seine Firma den Bauauftrag bekommen würde.  
*Continuation*: Die Bestechung des Unternehmers<sub>GenS</sub>/Beamten<sub>GenO</sub> entschied den Wettbewerb zu Gunsten der Firma.
58. *Context*: Der Verein bot dem Schiedsrichter viel Geld für einen Sieg im nächsten Spiel.  
*Continuation*: Die Bestechung des Vereins<sub>GenS</sub>/Schiedsrichters<sub>GenO</sub> veranlasste den Staatsanwalt zu einer Anklage.

#### **Bestrafung**

59. *Context*: Der Trainer hetzte den Sportler nach zwei Ermahnungen einige Male um den Platz.  
*Continuation*: Die Bestrafung des Trainers<sub>GenS</sub>/Sportlers<sub>GenO</sub> amüsierte das Team für einige Minuten.
60. *Context*: Der Lehrer gab dem Schüler nach seinen störenden Kommentaren eine zusätzliche Aufgabe.  
*Continuation*: Die Bestrafung des Lehrers<sub>GenS</sub>/Schülers<sub>GenO</sub> veranlasste die Mitschüler zum Aufpassen.

#### **Betreuung**

61. *Context*: Der Assistent unterstützte den Studenten während seiner Diplomarbeit.  
*Continuation*: Die Betreuung des Assistenten<sub>GenS</sub>/Studenten<sub>GenO</sub> hinterließ einen guten Eindruck bei den Kollegen.
62. *Context*: Der Enkel kümmerte sich lange um den kranken Großvater.  
*Continuation*: Die Betreuung des Enkels<sub>GenS</sub>/Großvaters<sub>GenO</sub> verbesserte das Verhältnis zur Großmutter.

#### **Bevormundung**

63. *Context*: Der Vater entschied, welche Ausbildung der Sohn machen sollte.  
*Continuation*: Die Bevormundung des Vaters<sub>GenS</sub>/Sohnes<sub>GenO</sub> belastete die Familie über eine lange Zeit.
64. *Context*: Der Arzt verweigerte dem Patienten die übliche Behandlung.  
*Continuation*: Die Bevormundung des Arztes<sub>GenS</sub>/Patienten<sub>GenO</sub> hatte ein Nachspiel vor Gericht.

#### **Bewachung**

65. *Context*: Der Leibwächter ließ den Rockstar während der gesamten Tournee nicht aus den Augen.  
*Continuation*: Die Bewachung des Leibwächters<sub>GenS</sub>/Rockstars<sub>GenO</sub> beruhigte den Manager während der Tournee.

66. *Context:* Der Polizist beschützte den Zeugen bis zum Ende des Prozesses.  
*Continuation:* Die Bewachung des Polizisten<sub>GenS</sub>/Zeugen<sub>GenO</sub> ermöglichte die Aussage gegen die Angeklagten.

#### **Bewunderung**

67. *Context:* Der Junge vergötterte den Fußballer über viele Jahre.  
*Continuation:* Die Bewunderung des Jungen<sub>GenS</sub>/Fußballers<sub>GenO</sub> nervte den Bruder nach einiger Zeit.
68. *Context:* Der Sohn himmelte den Star Tag und Nacht an.  
*Continuation:* Die Bewunderung des Sohnes<sub>GenS</sub>/Stars<sub>GenO</sub> beunruhigte den Vater nach einiger Zeit.

#### **Diskriminierung**

69. *Context:* Der Lehrer benachteiligte einen Schüler konstant bei der Benotung.  
*Continuation:* Die Diskriminierung des Lehrers<sub>GenS</sub>/Schülers<sub>GenO</sub> zwang die Eltern zu einer Beschwerde.
70. *Context:* Der Beamte benachteiligte den älteren Bewerber bei der Vergabe des Postens.  
*Continuation:* Die Diskriminierung des Beamten<sub>GenS</sub>/Bewerbers<sub>GenO</sub> entsetzte die Kollegen im Ministerium.

#### **Durchleuchtung**

71. *Context:* Der Ermittler überprüfte den Verdächtigen wochenlang von Kopf bis Fuß.  
*Continuation:* Die Durchleuchtung des Ermittlers<sub>GenS</sub>/Verdächtigen<sub>GenO</sub> erbrachte den Beweis für die Verurteilung.
72. *Context:* Der Staatsanwalt überprüfte den verdächtigen Politiker wieder und wieder.  
*Continuation:* Die Durchleuchtung des Staatsanwalts<sub>GenS</sub>/Politikers<sub>GenO</sub> beschäftigte die Kollegen in einer Parteisitzung.

#### **Durchsuchung**

73. *Context:* Der Kommissar überprüfte den Räuber bei der Festnahme auf Waffen.  
*Continuation:* Die Durchsuchung des Kommissars<sub>GenS</sub>/Räubers<sub>GenO</sub> lockte einige Schaulustige in die Nähe.
74. *Context:* Der Zöllner kontrollierte bei der Einreise, ob der Fahrer Drogen bei sich hatte.  
*Continuation:* Die Durchsuchung des Zöllners<sub>GenS</sub>/Fahrers<sub>GenO</sub> verzögerte die Einreise um einige Stunden.

#### **Einbürgerung**

75. *Context:* Der Minister sorgte dafür, dass der Sportler sofort die deutsche Staatsbürgerschaft erhielt.  
*Continuation:* Die Einbürgerung des Ministers<sub>GenS</sub>/Sportlers<sub>GenO</sub> erntete einigen Spott in der internationalen Presse.
76. *Context:* Der Beamte erkannte dem Künstler nach vielen Jahren im Land die Staatsbürgerschaft zu.  
*Continuation:* Die Einbürgerung des Beamten<sub>GenS</sub>/Künstlers<sub>GenO</sub> erleichterte das Arbeiten in der neuen Heimat.

#### **Einlieferung**

77. *Context:* Der Sanitäter fuhr das Opfer in das nächste Krankenhaus.  
*Continuation:* Die Einlieferung des Sanitäters<sub>GenS</sub>/Opfers<sub>GenO</sub> versetzte die Notaufnahme in Alarm.
78. *Context:* Der Arzt brachte das Kind mit dem Hubschrauber in das Krankenhaus.  
*Continuation:* Die Einlieferung des Arztes<sub>GenS</sub>/Kindes<sub>GenO</sub> bedeutete die Rettung in letzter Sekunde.

#### **Einschüchterung**

79. *Context:* Der Minister drohte dem Reporter mit einer Anzeige wegen Verleumdung.  
*Continuation:* Die Einschüchterung des Ministers<sub>GenS</sub>/Reporters<sub>GenO</sub> veranlasste die Zeitung zu einer Stellungnahme.
80. *Context:* Der Gangster drohte dem Zeugen vor dem Prozess.  
*Continuation:* Die Einschüchterung des Gangsters<sub>GenS</sub>/Zeugen<sub>GenO</sub> zeigte keine Wirkung während der Aussage.

### Enteignung

81. *Context:* Der Richter verteilte das Vermögen des Millionärs an die Gläubiger.  
*Continuation:* Die Enteignung des Richters<sub>GenS</sub>/Millionärs<sub>GenO</sub> verhalf den Gläubigern zu ihrem Geld.
82. *Context:* Der Diktator nahm dem Adel nach der Revolution den Grundbesitz weg.  
*Continuation:* Die Enteignung des Diktators<sub>GenS</sub>/Adels<sub>GenO</sub> veränderte die Verhältnisse in dem Land radikal.

### Enterbung

83. *Context:* Der Vater strich den Sohn aus dem Testament.  
*Continuation:* Die Enterbung des Vaters<sub>GenS</sub>/Sohnes<sub>GenO</sub> stürzte die Familie in einen Streit.
84. *Context:* Der Großvater beauftragte den Anwalt, den Enkel aus dem Testament zu nehmen.  
*Continuation:* Die Enterbung des Großvaters<sub>GenS</sub>/Enkels<sub>GenO</sub> überraschte die Familie nach den Ereignissen nicht besonders.

### Entführung

85. *Context:* Der Täter verschleppte den Jungen in ein abgelegenes Versteck.  
*Continuation:* Die Entführung des Täters<sub>GenS</sub>/Jungen<sub>GenO</sub> versetzte die Familie in große Angst.
86. *Context:* Der Terrorist hielt den Touristen bis zur Lösegeldzahlung gefangen.  
*Continuation:* Die Entführung des Terroristen<sub>GenS</sub>/Touristen<sub>GenO</sub> hielt die Angehörigen in Atem.

### Entgiftung

87. *Context:* Der Doktor betreute den Sohn während des erfolgreichen Drogenentzugs.  
*Continuation:* Die Entgiftung des Doktors<sub>GenS</sub>/Sohnes<sub>GenO</sub> erlöste die Eltern nach Jahren der Angst.
88. *Context:* Der Arzt begleitete den Süchtigen während der Entwöhnungsphase des Entzugs.  
*Continuation:* Die Entgiftung des Arztes<sub>GenS</sub>/Süchtigen<sub>GenO</sub> legte die Basis für eine erfolgreiche Therapie.

### Enthauptung

89. *Context:* Der Entführer tötete den Gefangenen vor laufender Kamera.  
*Continuation:* Die Enthauptung des Entführers<sub>GenS</sub>/Gefangenen<sub>GenO</sub> schockierte die Ermittler beim Ansehen des Videos.
90. *Context:* Im Film schlug der Ritter dem Zauberer den Kopf ab.  
*Continuation:* Die Enthauptung des Ritters<sub>GenS</sub>/Zauberers<sub>GenO</sub> schockierte einige Zuschauer im Kino.

### Enthebung

91. *Context:* Der Minister suspendierte den General nachdem die Vorwürfe an die Öffentlichkeit kamen.  
*Continuation:* Die Enthebung des Ministers<sub>GenS</sub>/Generals<sub>GenO</sub> veranlasste die Presse zu genauen Nachforschungen.
92. *Context:* Der Vorstand der Partei entzog dem Abgeordneten nach dem Bekanntwerden der Vorwürfe alle Funktionen.  
*Continuation:* Die Enthebung des Vorstands<sub>GenS</sub>/Abgeordneten<sub>GenO</sub> fand viel Unterstützung in der Partei.

### Entlarvung

93. *Context:* Der Detektiv meldete den Betrüger bei der Versicherung.  
*Continuation:* Die Entlarvung des Detektivs<sub>GenS</sub>/Betrügers<sub>GenO</sub> bestätigte die Versicherung in ihrem Verdacht.
94. *Context:* Der Informant bestätigte, dass der verdächtige Politiker ein Spion war.  
*Continuation:* Die Entlarvung des Informanten<sub>GenS</sub>/Spions<sub>GenO</sub> zwang den Kanzler zum Rücktritt.

### Entlassung

95. *Context:* Der Chef kündigte dem Arbeiter nach zwölf Jahren in der Firma.  
*Continuation:* Die Entlassung des Chefs<sub>GenS</sub>/Arbeiters<sub>GenO</sub> brachte die Familie in eine schwierige Situation.

96. *Context*: Der Chirurg schickte den Jungen fünf Wochen nach der Operation nach Hause.  
*Continuation*: Die Entlassung des Chirurgen<sub>GenS</sub>/Jungen<sub>GenO</sub> freute das Personal auf der Station.

#### Entlastung

97. *Context*: Der Praktikant half dem Angestellten drei Monate lang mit seiner Arbeit.  
*Continuation*: Die Entlastung des Praktikanten<sub>GenS</sub>/Angestellten<sub>GenO</sub> erleichterte die Arbeit während dieser Zeit sehr.
98. *Context*: Der Assistent nahm dem Chef viel Arbeit ab.  
*Continuation*: Die Entlastung des Assistenten<sub>GenS</sub>/Chefs<sub>GenO</sub> erleichterte die Arbeit an dem neuen Projekt.

#### Entmachtung

99. *Context*: Der General stürzte den Präsidenten kurz vor der Wahl.  
*Continuation*: Die Entmachtung des Generals<sub>GenS</sub>/Präsidenten<sub>GenO</sub> stürzte das Land in totales Chaos.
100. *Context*: Der Diktator entzog dem Parlament nach dem Putsch alle Kompetenzen.  
*Continuation*: Die Entmachtung des Diktators<sub>GenS</sub>/Parlaments<sub>GenO</sub> bedeutete das Ende für die Demokratie.

#### Entwaffnung

101. *Context*: Der Räuber nahm dem Wächter die Pistole weg.  
*Continuation*: Die Entwaffnung des Räubers<sub>GenS</sub>/Wächters<sub>GenO</sub> verursachte plötzliche Panik unter den Kunden der Bank.
102. *Context*: Der mutige Kunde schlug dem Räuber die Pistole aus der Hand.  
*Continuation*: Die Entwaffnung des Kunden<sub>GenS</sub>/Räubers<sub>GenO</sub> zwang den Komplizen zu einer schnellen Flucht.

#### Ergreifung

103. *Context*: Der Polizist schnappte den Mörder nach langen Ermittlungen.  
*Continuation*: Die Ergreifung des Polizisten<sub>GenS</sub>/Mörders<sub>GenO</sub> ermöglichte eine Verurteilung nach Jahren voller Rückschläge.
104. *Context*: Der Agent überwältigte den Terroristen nach einem kurzen Kampf.  
*Continuation*: Die Ergreifung des Agenten<sub>GenS</sub>/Terroristen<sub>GenO</sub> inspirierte einen Regisseur zu einem Film.

#### Ermahnung

105. *Context*: Der Schiedsrichter warnte den Verteidiger nach dem ersten Foul.  
*Continuation*: Die Ermahnung des Schiedsrichters<sub>GenS</sub>/Verteidigers<sub>GenO</sub> brachte den Trainer in Rage.
106. *Context*: Der Vater drohte dem Jungen mit Hausarrest, wenn er sein Zimmer nicht aufräumen würde.  
*Continuation*: Die Ermahnung des Vaters<sub>GenS</sub>/Jungen<sub>GenO</sub> brachte den Bruder zum Grinsen.

#### Ermordung

107. *Context*: Der Attentäter tötete den Politiker während einer Wahlveranstaltung.  
*Continuation*: Die Ermordung des Attentäters<sub>GenS</sub>/Politikers<sub>GenO</sub> verzögerte die Wahl um einige Monate.
108. *Context*: Der Räuber erschoss den Wächter während seiner Flucht aus der Bank.  
*Continuation*: Die Ermordung des Räubers<sub>GenS</sub>/Wächters<sub>GenO</sub> schockierte die Angestellten in der Bank.

#### Ermutigung

109. *Context*: Der Chef munterte den Arbeiter nach der Panne wieder auf.  
*Continuation*: Die Ermutigung des Chefs<sub>GenS</sub>/Arbeiters<sub>GenO</sub> überraschte die Kollegen in der Firma.
110. *Context*: Der Trainer richtete das Team nach der knappen Niederlage wieder etwas auf.  
*Continuation*: Die Ermutigung des Trainers<sub>GenS</sub>/Teams<sub>GenO</sub> machte die Niederlage für alle etwas erträglicher.

#### Erpressung



111. *Context:* Der Diener drohte dem Prinzen mit der Veröffentlichung von peinlichen Fotos.  
*Continuation:* Die Erpressung des Dieners<sub>GenS</sub>/Prinzen<sub>GenO</sub> setzte das Königshaus unter großen Druck.
112. *Context:* Der ehemalige Assistent drohte dem Manager damit, belastende Daten weiterzuleiten.  
*Continuation:* Die Erpressung des Assistenten<sub>GenS</sub>/Managers<sub>GenO</sub> brachte die Konzernführung in eine schwierige Situation.

#### **Förderung**

113. *Context:* Der Millionär unterstützte den Künstler mit einer beträchtlichen Summe.  
*Continuation:* Die Förderung des Millionärs<sub>GenS</sub>/Künstlers<sub>GenO</sub> ermöglichte eine Arbeit ohne großen Druck.
114. *Context:* Der Konzern garantierte dem Verein aus der zweiten Liga Zahlungen über zwei Jahre.  
*Continuation:* Die Förderung des Konzerns<sub>GenS</sub>/Vereins<sub>GenO</sub> ermöglichte den Aufstieg in die erste Liga.

#### **Freilassung**

115. *Context:* Der umstellte Täter ließ das Opfer nach Verhandlungen mit der Polizei gehen.  
*Continuation:* Die Freilassung des Täters<sub>GenS</sub>/Opfers<sub>GenO</sub> erlöste die Polizisten nach langem Warten.
116. *Context:* Der Kommissar schickte den Verdächtigen nach mehreren Verhören nach Hause.  
*Continuation:* Die Freilassung des Kommissars<sub>GenS</sub>/Verdächtigen<sub>GenO</sub> warf die Ermittlungen an den Anfang zurück.

#### **Heilung**

117. *Context:* Der Doktor befreite den Sportler von seinen Beschwerden.  
*Continuation:* Die Heilung des Doktors<sub>GenS</sub>/Sportlers<sub>GenO</sub> ermöglichte den Sieg bei der Weltmeisterschaft.
118. *Context:* Der Arzt erlöste den Jungen mit seiner neuen Methode von seiner Krankheit.  
*Continuation:* Die Heilung des Arztes<sub>GenS</sub>/Patienten<sub>GenO</sub> erlöste die Familie nach Monaten voller Angst.

#### **Inhaftierung**

119. *Context:* Der Inspektor steckte den Verdächtigen in Untersuchungshaft.  
*Continuation:* Die Inhaftierung des Inspektors<sub>GenS</sub>/Verdächtigen<sub>GenO</sub> brachte das Verfahren ins Rollen.
120. *Context:* Der Polizist steckte den Betrunkenen zum Ausnüchtern in die Zelle.  
*Continuation:* Die Inhaftierung des Polizisten<sub>GenS</sub>/Betrunkenen<sub>GenO</sub> beendete den Dienst für diese Nacht.

#### **Manipulierung**

121. *Context:* Der Gangster versprach dem Geschworenen im Fall eines Freispruchs eine hohe Summe.  
*Continuation:* Die Manipulierung des Gangsters<sub>GenS</sub>/Geschworenen<sub>GenO</sub> hatte keine Auswirkung auf das Urteil.
122. *Context:* Der Pate warnte den Zeugen vor einer Aussage gegen ihn.  
*Continuation:* Die Manipulierung des Paten<sub>GenS</sub>/Zeugen<sub>GenO</sub> gab dem Prozess eine unerwartete Wende.

#### **Rehabilitierung**

123. *Context:* Der Richter hob das Urteil gegen den Mann nach 19 Jahren auf.  
*Continuation:* Die Rehabilitierung des Richters<sub>GenS</sub>/Mannes<sub>GenO</sub> rührte den Anwalt zu Tränen.
124. *Context:* Der Detektiv bewies die Unschuld des Angeklagten vor Gericht.  
*Continuation:* Die Rehabilitierung des Detektivs<sub>GenS</sub>/Angeklagten<sub>GenO</sub> ermöglichte der Familie, in Ruhe zu leben.

#### **Rettung**

125. *Context:* Der Passant holte den Fahrer in letzter Minute aus dem Autowrack.  
*Continuation:* Die Rettung des Passanten<sub>GenS</sub>/Fahrers<sub>GenO</sub> beherrschte die Titelblätter in den lokalen Zeitungen.

126. *Context*: Der Hund fand das Opfer der Lawine noch rechtzeitig unter dem Schnee.  
*Continuation*: Die Rettung des Hundes<sub>GenS</sub>/Opfers<sub>GenO</sub> erlöste die Eltern nach Stunden voller Ungewissheit.

#### Schulung

127. *Context*: Der Experte erklärte dem Arbeiter in dem Seminar die neue Maschine.  
*Continuation*: Die Schulung des Experten<sub>GenS</sub>/Arbeiter<sub>GenO</sub> erleichterte die Arbeit an der Maschine um Einiges.
128. *Context*: Der Chef wollte das neue Team persönlich in die Software einführen.  
*Continuation*: Die Schulung des Chefs<sub>GenS</sub>/Teams<sub>GenO</sub> lieferte die Grundlage für die weitere Arbeit.

#### Täuschung

129. *Context*: Der Manager schickte dem Richter gefälschte Unterlagen.  
*Continuation*: Die Täuschung des Managers<sub>GenS</sub>/Richters<sub>GenO</sub> hatte ein Nachspiel für den ganzen Konzern.
130. *Context*: Der Sportler übergab dem Kontrolleur eine manipulierte Dopingprobe.  
*Continuation*: Die Täuschung des Sportlers<sub>GenS</sub>/Kontrollieurs<sub>GenO</sub> schockierte die Fans als sie aufgedeckt wurde.

#### Tötung

131. *Context*: Der Wilderer erschoss den Löwen mitten im Nationalpark.  
*Continuation*: Die Tötung des Wilderers<sub>GenS</sub>/Löwen<sub>GenO</sub> überraschte die Ranger nach Jahren der Ruhe.
132. *Context*: Der Direktor des Zoos ließ den jungen Tiger nach dem Vorfall einschläfern.  
*Continuation*: Die Tötung des Direktors<sub>GenS</sub>/Tigers<sub>GenO</sub> ertete viel Kritik von den Medien.

#### Überprüfung

133. *Context*: Der Beamte befragte den Touristen vor der Einreise in das Land.  
*Continuation*: Die Überprüfung des Beamten<sub>GenS</sub>/Touristen<sub>GenO</sub> brachte keinen Anlass zur Sorge.
134. *Context*: Der Detektiv durchleuchtete den Verdächtigen von Kopf bis Fuß.  
*Continuation*: Die Überprüfung des Detektivs<sub>GenS</sub>/Verdächtigen<sub>GenO</sub> förderte einige Beweise zu Tage.

#### Überstellung

135. *Context*: Der Wärter begleitete den Häftling in ein neues Gefängnis.  
*Continuation*: Die Überstellung des Wärters<sub>GenS</sub>/Häftlings<sub>GenO</sub> ermöglichte eine Verwahrung ohne Sicherheitsrisiko.
136. *Context*: Der Pfleger brachte den Patienten von der Intensivstation in ein normales Zimmer.  
*Continuation*: Die Überstellung des Pflegers<sub>GenS</sub>/Patienten<sub>GenO</sub> erfüllte die Familie mit großer Hoffnung.

#### Überwachung

137. *Context*: Der Detektiv beobachtete den Verdächtigen über mehrere Monate.  
*Continuation*: Die Überwachung des Detektivs<sub>GenS</sub>/Verdächtigen<sub>GenO</sub> brachte den Auftraggeber um viel Geld.
138. *Context*: Der Agent folgte dem Diplomaten über mehrere Wochen.  
*Continuation*: Die Überwachung des Agenten<sub>GenS</sub>/Diplomaten<sub>GenO</sub> erbrachte keine Hinweise auf mögliche Spionage.

#### Umarmung

139. *Context*: Der Vater schloss den Sohn nach 20 Jahren Streit in seine Arme.  
*Continuation*: Die Umarmung des Vaters<sub>GenS</sub>/Sohnes<sub>GenO</sub> rührte die Mutter zu Tränen.
140. *Context*: Der Sänger fiel dem Dirigent nach dem Konzert um den Hals.  
*Continuation*: Die Umarmung des Sängers<sub>GenS</sub>/Dirigenten<sub>GenO</sub> veranlasste das Publikum zu Jubelrufen.

#### Unterdrückung

141. *Context:* Der Diktator beutete das Volk brutal und gnadenlos aus.  
*Continuation:* Die Unterdrückung des Diktators<sub>GenS</sub>/Volkes<sub>GenO</sub> veranlasste die UNO zu einer gemeinsamen Resolution.
142. *Context:* Der Tyrann bekämpfte den Widerstand mit brutalen Mitteln.  
*Continuation:* Die Unterdrückung des Tyrannen<sub>GenS</sub>/Widerstandes<sub>GenO</sub> forderte viele Opfer unter den Rebellen.

#### Unterstützung

143. *Context:* Der Konzern förderte den Sportler mit einer großen Summe.  
*Continuation:* Die Unterstützung des Konzerns<sub>GenS</sub>/Sportlers<sub>GenO</sub> hatte positive Auswirkungen auf das Image der Firma.
144. *Context:* Der Vorstand deckte dem Manager nach dem Skandal den Rücken.  
*Continuation:* Die Unterstützung des Vorstands<sub>GenS</sub>/Managers<sub>GenO</sub> verursachte massive Proteste von Politikern und Öffentlichkeit.

#### Untersuchung

145. *Context:* Der Doktor überprüfte den Mann auf mögliche neurologische Störungen.  
*Continuation:* Die Untersuchung des Doktors<sub>GenS</sub>/Mannes<sub>GenO</sub> brachte keinen Hinweis auf eine Erkrankung.
146. *Context:* Der Prüfer des Finanzamts kontrollierte den Millionär.  
*Continuation:* Die Untersuchung des Prüfers<sub>GenS</sub>/Millionärs<sub>GenO</sub> ergab einige Hinweise auf Unregelmäßigkeiten.

#### Verbannung

147. *Context:* Der König schickte den aufständischen General ins Exil.  
*Continuation:* Die Verbannung des Königs<sub>GenS</sub>/Generals<sub>GenO</sub> beendete den Aufstand gegen die Monarchie.
148. *Context:* Der Präsident verwies den kritischen Journalisten aus dem Land.  
*Continuation:* Die Verbannung des Präsidenten<sub>GenS</sub>/Journalisten<sub>GenO</sub> bedeutete einen Schlag gegen die Pressefreiheit.

#### Verhaftung

149. *Context:* Der Polizist schnappte den Dieb nach einer wilden Verfolgungsjagd.  
*Continuation:* Die Verhaftung des Polizisten<sub>GenS</sub>/Diebs<sub>GenO</sub> lockte einige Anwohner aus ihren Häusern.
150. *Context:* Der Kommissar führte den bekannten Sänger in Handschellen ab.  
*Continuation:* Die Verhaftung des Kommissars<sub>GenS</sub>/Sängers<sub>GenO</sub> versetzte die Fans in große Aufregung.

#### Verlegung

151. *Context:* Der Pfleger brachte den Patienten von der Intensivstation in die Chirurgie.  
*Continuation:* Die Verlegung des Pflegers<sub>GenS</sub>/Patienten<sub>GenO</sub> überraschte den Arzt auf der Intensivstation.
152. *Context:* Der Direktor des Gefängnisses ließ den Gefangenen in einen anderen Trakt bringen.  
*Continuation:* Die Verlegung des Direktors<sub>GenS</sub>/Gefangenen<sub>GenO</sub> trennte die Bande von ihrem Chef.

#### Vernehmung

153. *Context:* Der Kommissar befragte den verhafteten Kriminellen über die Organisation.  
*Continuation:* Die Vernehmung des Kommissars<sub>GenS</sub>/Kriminellen<sub>GenO</sub> brachte viele Informationen für den Staatsanwalt.
154. *Context:* Der Staatsanwalt befragte den Kronzeugen im Zeugenstand.  
*Continuation:* Die Vernehmung des Staatsanwalts<sub>GenS</sub>/Kronzeugen<sub>GenO</sub> überzeugte die Jury von der Schuld des Angeklagten.

#### Verschleppung

155. *Context:* Der Spion entführte den Forscher aus seinem Heimatland.  
*Continuation:* Die Verschleppung des Spions<sub>GenS</sub>/Forschers<sub>GenO</sub> verschärfte die Spannungen zwischen den Ländern.
156. *Context:* Der Vater brachte das Kind ohne Wissen der Mutter in sein Heimatland.  
*Continuation:* Die Verschleppung des Vaters<sub>GenS</sub>/Kindes<sub>GenO</sub> brachte die Mutter zur Verzweiflung.

#### **Versetzung**

157. *Context:* Der General wies den Soldaten einer neuen Einheit zu.  
*Continuation:* Die Versetzung des Generals<sub>GenS</sub>/Soldaten<sub>GenO</sub> erleichterte die Kollegen in der Kompanie.
158. *Context:* Der Chef schickte den Mitarbeiter zum Aufbau des neuen Büros nach Asien.  
*Continuation:* Die Versetzung des Chefs<sub>GenS</sub>/Mitarbeiters<sub>GenO</sub> stellte die Familie vor große Probleme.

#### **Versorgung**

159. *Context:* Der Sponsor sorgte während des zweiwöchigen Radrennens dafür, dass der Fahrer gut betreut wurde.  
*Continuation:* Die Versorgung des Sponsors<sub>GenS</sub>/Fahrers<sub>GenO</sub> erleichterte die Arbeit für das gesamte Team.
160. *Context:* Der Koch stellte während der Expedition sicher, dass das Team gut verpflegt wurde.  
*Continuation:* Die Versorgung des Kochs<sub>GenS</sub>/Teams<sub>GenO</sub> verbesserte die Stimmung während des Abenteuers.

#### **Verteidigung**

161. *Context:* Der Anwalt vertrat den Manager in einem schwierigen Prozess.  
*Continuation:* Die Verteidigung des Anwalts<sub>GenS</sub>/Managers<sub>GenO</sub> erforderte eine Strategie mit viel Risiko.
162. *Context:* Der Neue hielt den Stürmer das ganze Spiel über unter Kontrolle.  
*Continuation:* Die Verteidigung des Neuen<sub>GenS</sub>/Stürmers<sub>GenO</sub> beeindruckte den Trainer wegen ihrer Effektivität.

#### **Vertreibung**

163. *Context:* Der Besitzer zwang den Mieter, aus der Wohnung auszuziehen.  
*Continuation:* Die Vertreibung des Besitzers<sub>GenS</sub>/Mieters<sub>GenO</sub> veranlasste den Anwalt zu einer Klage.
164. *Context:* Der General jagte den König nach dem Putsch aus dem Land.  
*Continuation:* Die Vertreibung des Generals<sub>GenS</sub>/Königs<sub>GenO</sub> beunruhigte die Regierungen in den Nachbarländern.

#### **Verurteilung**

165. *Context:* Der Richter schickte den Mann für zehn Jahre ins Gefängnis.  
*Continuation:* Die Verurteilung des Richters<sub>GenS</sub>/Mannes<sub>GenO</sub> überraschte die Zuseher im Gerichtssaal.
166. *Context:* Das Gremium sperrte den Sportler nach den positiven Dopingtests für zwei Jahre.  
*Continuation:* Die Verurteilung des Gremiums<sub>GenS</sub>/Sportlers<sub>GenO</sub> beendete die Karriere nach zehn Jahren im Profisport.

#### **Würdigung**

167. *Context:* Der Vorgesetzte lobte den Angestellten in der Besprechung in den Himmel.  
*Continuation:* Die Würdigung des Vorgesetzten<sub>GenS</sub>/Angestellten<sub>GenO</sub> beeindruckte die Kollegen in der Firma.
168. *Context:* Der Minister betonte in seiner Rede die Verdienste des Künstlers.  
*Continuation:* Die Würdigung des Ministers<sub>GenS</sub>/Künstlers<sub>GenO</sub> beeindruckte viele Zuseher im Saal.

#### **Zurückweisung**

169. *Context:* Der Sohn vergab seinem Vater auch nach jahrelangem Streit nicht.  
*Continuation:* Die Zurückweisung des Sohnes<sub>GenS</sub>/Vaters<sub>GenO</sub> stellte die Familie auf eine harte Probe.
170. *Context:* Der Großvater akzeptierte seinen unehelichen Enkel nie.  
*Continuation:* Die Zurückweisung des Großvaters<sub>GenS</sub>/Enkels<sub>GenO</sub> belastete die Familie über viele Jahre hinweg.

## Appendix C

# Supplementary materials for the self-paced reading study

### Supplementary analyses

#### Rating data

The following tables summarise the additional cumulative link mixed models (CLMMs) fitted on the rating data of the SPR experiment using function `clmm` from the *ordinal* package (Christensen, 2013b). Table C.1 presents the estimates for the full data set, while tables C.2 and C.3 summarise the CLMMs fitted on the GenO and GenS subsets, respectively.

##### Random effects

Groups	Name	Variance	SD	Corr.		
Story	(Intercept)	0.31	0.56			
	c.Linking	1.53	1.24	-0.05		
Subject	(Intercept)	0.29	0.54			
	c.Linking	2.39	1.55	-0.15		
	c.AI	0.02	0.15	0.07	0.28	

##### Fixed effects

	Estimate	SE	Lower	Upper	z-value	Pr(> z )	
c.Linking	-3.05	0.37	-3.77	-2.33	-8.27	<0.001	***
c.AI	-0.21	0.06	-0.33	-0.10	-3.79	<0.001	***
c.Linking:c.AI	-1.05	0.10	-1.24	-0.86	-10.89	<0.001	***

**Table C.1:** Summary tables for the cumulative link mixed model (CLMM) for the acceptability rating data of the SPR experiment.

Random effects

Groups	Name	Variance	SD	Corr.
Story	(Intercept)	0.77	0.88	
Subject	(Intercept)	1.07	1.04	
	c.AI	0.01	0.12	0.43

Fixed effects

	Estimate	SE	Lower	Upper	z-value	Pr(> z )
c.AI	0.33	0.07	0.18	0.47	4.45	<0.001 ***

**Table C.2:** CLMM results for the GenO subset of the SPR acceptability data.

Random effects

Groups	Name	Variance	SD	Corr.
Story	(Intercept)	0.69	0.83	
Subject	(Intercept)	0.76	0.87	
	c.AI	0.06	0.24	0.29

Fixed effects

	Estimate	SE	Lower	Upper	z-value	Pr(> z )
c.AI	-0.76	0.09	-0.93	-0.59	-8.83	<0.001 ***

**Table C.3:** CLMM results for the GenS subset of the SPR acceptability data.

## Stimulus sentences

The following pages list the stimulus sentences used in the experiment, alphabetically ordered by the relevant *-ung*-nominalisation and numbered consecutively by 'story'. The first sentence represents the context sentence, the second the continuation sentence; GenO and GenS subscripts indicate the classifications of the genitives presented in the continuation sentences.

### Ablehnung

1. *Context:* Der Direktor weigerte sich, den Forscher einzustellen.  
*Continuation:* Die Ablehnung des Direktors<sub>GenS</sub>/Forschers<sub>GenO</sub> enttäuschte die Kollegen im Labor.
2. *Context:* Der Papst verweigerte dem Kardinal das hohe Amt im Vatikan.  
*Continuation:* Viele Journalisten bezeichneten die Ablehnung des Papstes<sub>GenS</sub>/Kardinals<sub>GenO</sub> als falsches Signal.

### Abschirmung

3. *Context:* Der Anwalt hielt sämtliche Reporter von dem Opfer fern.  
*Continuation:* Die Abschirmung des Anwalts<sub>GenS</sub>/Opfers<sub>GenO</sub> ermöglichte die Vorbereitung auf den Prozess.
4. *Context:* Der Bodyguard ließ während des gesamten Termins keine Fans an den Sportler heran.  
*Continuation:* Die Fans hielten die Abschirmung des Bodyguards<sub>GenS</sub>/Sportlers<sub>GenO</sub> für übertrieben.

### Absetzung

5. *Context:* Der Kanzler entließ seinen Minister, nachdem der Skandal aufgefliegen war.  
*Continuation:* Die Absetzung des Kanzlers<sub>GenS</sub>/Ministers<sub>GenO</sub> rettete die Regierung vor einer Krise.
6. *Context:* Der Rektor enthob den Professor nach dem Skandal von seinem Posten.  
*Continuation:* Die Studenten unterstützten die Absetzung des Rektors<sub>GenS</sub>/Professors<sub>GenO</sub> zum größten Teil.

### Abwerbung

7. *Context:* Der Chef des Konzerns brachte den Forscher dazu, in sein Unternehmen zu wechseln.  
*Continuation:* Die Abwerbung des Chefs<sub>GenS</sub>/Forschers<sub>GenO</sub> hatte ein Nachspiel vor Gericht.
8. *Context:* Der Manager des Fußballclubs kaufte den Star des Erzrivalen für seinen eigenen Verein.  
*Continuation:* Die Fans quittierten die Abwerbung des Managers<sub>GenS</sub>/Stars<sub>GenO</sub> im nächsten Spiel mit Pfeifkonzerten.

### Anfeindung

9. *Context:* Der Professor attackierte den Moderator während einer Diskussion verbal.  
*Continuation:* Die Anfeindung des Professors<sub>GenS</sub>/Moderators<sub>GenO</sub> irritierte die Zuhörer im Saal.
10. *Context:* Der Politiker beschimpfte den Reporter während der Aufzeichnung des Interviews.  
*Continuation:* Der Regisseur entfernte die Anfeindung des Politikers<sub>GenS</sub>/Reporters<sub>GenO</sub> vor der Ausstrahlung.

### Anhörung

11. *Context:* Der Vorstand befragte den Manager, um seine Sicht der Dinge zu hören.  
*Continuation:* Die Anhörung des Vorstands<sub>GenS</sub>/Managers<sub>GenO</sub> erbrachte keine Hinweise auf einen Betrug.
12. *Context:* Der Senat befragte den Minister zu den Vorgängen.  
*Continuation:* Ein Fernsehsender übertrug die Anhörung des Senats<sub>GenS</sub>/Ministers<sub>GenO</sub> in voller Länge.

### Aufmunterung

13. *Context:* Der Vater tröstete das Kind nach der verpatzten Klausur.  
*Continuation:* Die Aufmunterung des Vaters<sub>GenS</sub>/Kindes<sub>GenO</sub> rettete das Wochenende für die ganze Familie.
14. *Context:* Der Clown brachte den kleinen Patienten nach langer Zeit wieder zum Lachen.  
*Continuation:* Die Ärztin verfolgte die Aufmunterung des Clowns<sub>GenS</sub>/Patienten<sub>GenO</sub> aus dem Hintergrund.

### Ausbeutung

15. *Context:* Der Professor übertrug dem Assistenten jahrelang seine eigene Arbeit.  
*Continuation:* Die Ausbeutung des Professors<sub>GenS</sub>/Assistenten<sub>GenO</sub> verärgerte viele Kollegen am Institut.
16. *Context:* Der Wirt ließ seinen Kellner für zu wenig Geld viel zu viel arbeiten.  
*Continuation:* Die Kündigung beendete die Ausbeutung des Wirts<sub>GenS</sub>/Kellners<sub>GenO</sub> nach fast drei Jahren.

### Ausbildung

17. *Context:* Der Koch brachte dem Azubi im Lauf der Zeit viele neue Sachen bei.  
*Continuation:* Die Ausbildung des Kochs<sub>GenS</sub>/Azubis<sub>GenO</sub> beeindruckte den Chef des kleinen Restaurants.
18. *Context:* Der Meister vermittelte dem Lehrling sein ganzes Können mit viel Begeisterung.  
*Continuation:* Die Eltern lobten die Ausbildung des Meisters<sub>GenS</sub>/Lehrlings<sub>GenO</sub> bei jeder Gelegenheit.

### Auslieferung

19. *Context:* Der Senat schickte den geflohenen Diktator zurück in sein Land.  
*Continuation:* Die Auslieferung des Senats<sub>GenS</sub>/Diktators<sub>GenO</sub> verbesserte die Stimmung zwischen den Staaten.

20. *Context:* Der Richter überstellte den Verdächtigen nach langem Zögern in seine Heimat.  
*Continuation:* Die Behörden begrüßten die Auslieferung des Richters<sub>GenS</sub>/Verdächtigen<sub>GenO</sub> in einem Statement.

#### **Auswechslung**

21. *Context:* Der Kläger suchte sich kurz vor dem Prozess einen neuen Anwalt.  
*Continuation:* Die Auswechslung des Klägers<sub>GenS</sub>/Anwalts<sub>GenO</sub> verzögerte den Prozessbeginn um einige Wochen.
22. *Context:* Der Trainer nahm den Stürmer kurz nach der Halbzeitpause vom Feld.  
*Continuation:* Die Fans begleiteten die Auswechslung des Trainers<sub>GenS</sub>/Stürmers<sub>GenO</sub> mit einem Pfeifkonzert.

#### **Ausweisung**

23. *Context:* Der Beamte lehnte den Antrag des Flüchtlings auf Asyl ab.  
*Continuation:* Die Ausweisung des Beamten<sub>GenS</sub>/Flüchtlings<sub>GenO</sub> schockierte viele Leute wegen ihrer Härte.
24. *Context:* Der Richter entschied, den Asylanten in sein Heimatland zurückzuschicken.  
*Continuation:* Der Anwalt bekämpfte die Ausweisung des Richters<sub>GenS</sub>/Asylanten<sub>GenO</sub> in der Revision.

#### **Beförderung**

25. *Context:* Der Manager ernannte den Assistenten zum Projektleiter.  
*Continuation:* Die Beförderung des Managers<sub>GenS</sub>/Assistenten<sub>GenO</sub> freute die Kollegen im Büro.
26. *Context:* Der Minister ernannte den Beamten zum Botschafter.  
*Continuation:* Die Kollegen kritisierten die Beförderung des Ministers<sub>GenS</sub>/Beamten<sub>GenO</sub> intern.

#### **Befragung**

27. *Context:* Der Polizist verhörte den Zeugen zum Ablauf des Überfalls.  
*Continuation:* Die Befragung des Polizisten<sub>GenS</sub>/Zeugen<sub>GenO</sub> erbrachte einen Hinweis auf das Fluchtauto.
28. *Context:* Der Staatsanwalt nahm den Angeklagten ins Kreuzverhör.  
*Continuation:* Der Verteidiger unterbrach die Befragung des Staatsanwalts<sub>GenS</sub>/Angeklagten<sub>GenO</sub> mit mehreren Einsprüchen.

#### **Befreiung**

29. *Context:* Der Kommissar holte den verschleppten Jungen aus dem Keller des Entführers.  
*Continuation:* Die Befreiung des Kommissars<sub>GenS</sub>/Jungen<sub>GenO</sub> brachte den Fall in die internationalen Medien.
30. *Context:* Das Team der Spezialeinheit erlöste den Politiker aus der Geiselhaft.  
*Continuation:* Ein Offizier schilderte die Befreiung des Teams<sub>GenS</sub>/Politikers<sub>GenO</sub> in einem Interview.

#### **Begrüßung**

31. *Context:* Der Sänger richtete zu Beginn einige nette Worte an das Publikum.  
*Continuation:* Die Begrüßung des Sängers<sub>GenS</sub>/Publikums<sub>GenO</sub> eröffnete das Konzert in dem kleinen Club.
32. *Context:* Der Minister hieß den Athleten nach seinem Olympiasieg persönlich willkommen.  
*Continuation:* Der Reporter kommentierte die Begrüßung des Ministers<sub>GenS</sub>/Athleten<sub>GenO</sub> im Radio.

#### **Behandlung**

33. *Context:* Der Therapeut massierte den Sportler mit einer neuen Technik.  
*Continuation:* Die Behandlung des Therapeuten<sub>GenS</sub>/Sportlers<sub>GenO</sub> verringerte die Schmerzen nach wenigen Sitzungen.
34. *Context:* Der Hausarzt therapierte den Kranken mit einer alternativen Methode.  
*Continuation:* Die Kasse bezahlte die Behandlung des Hausarztes<sub>GenS</sub>/Kranken<sub>GenO</sub> zur Hälfte.

#### **Bekehrung**



35. *Context*: Der Priester überzeugte den Häftling davon, in die Kirche einzutreten.  
*Continuation*: Die Bekehrung des Priesters<sub>GenS</sub>/Häftlings<sub>GenO</sub> brachte den Direktor zum Staunen.
36. *Context*: Der Therapeut überzeugte den Raucher davon, mit dem Qualmen aufzuhören.  
*Continuation*: Die Familie bezeichnete die Bekehrung des Therapeuten<sub>GenS</sub>/Rachters<sub>GenO</sub> als ein kleines Wunder.

#### **Belagerung**

37. *Context*: Der Fan campierte monatelang vor dem Grundstück des Sängers.  
*Continuation*: Die Belagerung des Fans<sub>GenS</sub>/Sängers<sub>GenO</sub> zwang die Polizei nach einiger Zeit zum Handeln.
38. *Context*: Der Kaiser schnitt mit seiner Armee den Papst lange komplett von der Außenwelt ab.  
*Continuation*: Ein Kompromiss beendete die Belagerung des Kaisers<sub>GenS</sub>/Papstes<sub>GenO</sub> nach sechs Monaten.

#### **Belästigung**

39. *Context*: Ein Betrunkener provozierte den Wächter in der U-Bahn-Station.  
*Continuation*: Die Belästigung des Betrunkenen<sub>GenS</sub>/Wächters<sub>GenO</sub> veranlasste einen Fahrgast zum Eingreifen.
40. *Context*: Der Jugendliche bedrängte den Pensionisten auf offener Straße.  
*Continuation*: Ein Passant bemerkte die Belästigung des Jugendlichen<sub>GenS</sub>/Pensionisten<sub>GenO</sub> gerade noch rechtzeitig.

#### **Belehrung**

41. *Context*: Der Chef wies den nachlässigen Arbeiter auf seine Pflichten hin.  
*Continuation*: Die Belehrung des Chefs<sub>GenS</sub>/Arbeiters<sub>GenO</sub> verbesserte die Situation in den folgenden Wochen.
42. *Context*: Der Anwalt klärte den Angeklagten vor dem Prozess über seine Rechte auf.  
*Continuation*: Der Assistent unterbrach die Belehrung des Anwalts<sub>GenS</sub>/Angeklagten<sub>GenO</sub> wegen eines dringenden Telefonats.

#### **Beleidigung**

43. *Context*: Der Schüler beschimpfte den Lehrer in der Pause.  
*Continuation*: Die Beleidigung des Schülers<sub>GenS</sub>/Lehrers<sub>GenO</sub> schockierte die Mitschüler aus seiner Klasse.
44. *Context*: Der Fahrer zeigte dem Polizisten während einer Kontrolle den Mittelfinger.  
*Continuation*: Ein Kollege bestätigte die Beleidigung des Fahrers<sub>GenS</sub>/Polizisten<sub>GenO</sub> vor Gericht.

#### **Beratung**

45. *Context*: Der Experte half dem Vorstand des Unternehmens mit dem neuen Projekt.  
*Continuation*: Die Beratung des Experten<sub>GenS</sub>/Vorstands<sub>GenO</sub> kostete das Unternehmen ein kleines Vermögen.
46. *Context*: Der Fachmann erstellte dem Team vor dem Wahlkampf eine umfassende Analyse.  
*Continuation*: Die Partei bezahlte die Beratung des Fachmanns<sub>GenS</sub>/Teams<sub>GenO</sub> aus der Wahlkampfkasse.

#### **Bergung**

47. *Context*: Der Sanitäter zog den Fahrer vorsichtig aus dem Autowrack.  
*Continuation*: Die Bergung des Sanitäters<sub>GenS</sub>/Fahrers<sub>GenO</sub> ermöglichte dem Arzt eine schnelle Behandlung.
48. *Context*: Der Pilot des Hubschraubers zog den gekenterten Segler aus der stürmischen See.  
*Continuation*: Die Küstenwache meldete die Bergung des Piloten<sub>GenS</sub>/Seglers<sub>GenO</sub> an die Angehörigen.

#### **Beschattung**

49. *Context*: Der Agent beobachtete den Politiker über mehrere Wochen.  
*Continuation*: Die Beschattung des Agenten<sub>GenS</sub>/Politikers<sub>GenO</sub> brachte viele Informationen für die geplante Aktion.

50. *Context:* Der Geheimdienst ließ den Drogendealer während seines Aufenthalts nicht aus den Augen.  
*Continuation:* Der General koordinierte die Beschattung des Geheimdienstes<sub>GenS</sub>/Drogendealers<sub>GenO</sub> persönlich.

#### **Beschimpfung**

51. *Context:* Der Chef schockierte den Arbeiter mit seinen verletzenden Bemerkungen.  
*Continuation:* Die Beschimpfung des Chefs<sub>GenS</sub>/Arbeiters<sub>GenO</sub> verschlechterte die Stimmung noch weiter.
52. *Context:* Der Stürmer beleidigte den Tormann nach dem gehaltenen Elfmeter.  
*Continuation:* Der Schiedsrichter ahndete die Beschimpfung des Stürmers<sub>GenS</sub>/Tormanns<sub>GenO</sub> mit der roten Karte.

#### **Beseitigung**

53. *Context:* Der Agent vergiftete den Diktator beim Abendessen.  
*Continuation:* Die Beseitigung des Agenten<sub>GenS</sub>/Diktators<sub>GenO</sub> öffnete den Weg zu einem Machtwechsel.
54. *Context:* Der Manager der Firma trieb den Händler gezielt in den Konkurs.  
*Continuation:* Ein Journalist enthüllte die Beseitigung des Managers<sub>GenS</sub>/Händlers<sub>GenO</sub> in einem Artikel.

#### **Bespitzelung**

55. *Context:* Der Vorstand überprüfte die Telefonate des Managers.  
*Continuation:* Die Bespitzelung des Vorstands<sub>GenS</sub>/Managers<sub>GenO</sub> verursachte einen Skandal im Unternehmen.
56. *Context:* Der Spion hörte das Telefon des Ministers ab.  
*Continuation:* Ein Techniker enthüllte die Bespitzelung des Spions<sub>GenS</sub>/Ministers<sub>GenO</sub> nach einiger Zeit durch Zufall.

#### **Bestechung**

57. *Context:* Der Unternehmer versprach dem Beamten eine Menge Geld für den Bauauftrag.  
*Continuation:* Die Bestechung des Unternehmers<sub>GenS</sub>/Beamten<sub>GenO</sub> entschied den Wettbewerb zu Gunsten der Firma.
58. *Context:* Der Vorstand bezahlte den Betriebsrat für sein Schweigen.  
*Continuation:* Ein Kollege meldete die Bestechung des Vorstands<sub>GenS</sub>/Betriebsrats<sub>GenO</sub> an den Staatsanwalt.

#### **Bestrafung**

59. *Context:* Der Erzieher stellte den Jungen für eine halbe Stunde in die Ecke.  
*Continuation:* Die Bestrafung des Erziehers<sub>GenS</sub>/Jungen<sub>GenO</sub> veranlasste die Eltern zu einer Beschwerde.
60. *Context:* Der Richter verurteilte den Täter nur zu mehreren Wochen Sozialarbeit.  
*Continuation:* Der Anwalt bezeichnete die Bestrafung des Richters<sub>GenS</sub>/Täters<sub>GenO</sub> als viel zu milde.

#### **Betreuung**

61. *Context:* Der Pfleger beaufsichtigte den behinderten Jungen tagsüber.  
*Continuation:* Die Betreuung des Pflegers<sub>GenS</sub>/Jungen<sub>GenO</sub> erleichterte das Leben für die Eltern ungemein.
62. *Context:* Der Dolmetscher kümmerte sich während der gesamten Geschäftsreise um den Manager.  
*Continuation:* Der Gastgeber bezahlte die Betreuung des Dolmetschers<sub>GenS</sub>/Managers<sub>GenO</sub> aus der eigenen Tasche.

#### **Bevormundung**

63. *Context:* Der Arzt verweigerte dem Patienten die übliche Behandlung.  
*Continuation:* Die Bevormundung des Arztes<sub>GenS</sub>/Patienten<sub>GenO</sub> hatte ein Nachspiel vor Gericht.
64. *Context:* Der Vater entschied, welche Ausbildung der Sohn machen sollte.  
*Continuation:* Der Großvater kritisierte die Bevormundung des Vaters<sub>GenS</sub>/Sohnes<sub>GenO</sub> bei einem gemeinsamen Essen.

### **Bewachung**

65. *Context:* Der Polizist beschützte den Zeugen bis zum Ende des Prozesses.  
*Continuation:* Die Bewachung des Polizisten<sub>GenS</sub>/Zeugen<sub>GenO</sub> ließ die Familie in Ruhe schlafen.
66. *Context:* Der Agent schützte den Zeugen bis zum Prozess.  
*Continuation:* Der Staatsanwalt hielt die Bewachung des Agenten<sub>GenS</sub>/Zeugen<sub>GenO</sub> für eine notwendige Maßnahme.

### **Bewunderung**

67. *Context:* Der Enkel betete den Großvater an.  
*Continuation:* Die Bewunderung des Enkels<sub>GenS</sub>/Großvaters<sub>GenO</sub> machte den Vater manchmal etwas eifersüchtig.
68. *Context:* Das Mädchen reiste dem Popstar wochenlang hinterher.  
*Continuation:* Der Bruder hielt die Bewunderung des Mädchens<sub>GenS</sub>/Popstars<sub>GenO</sub> für leicht verrückt.

### **Diskriminierung**

69. *Context:* Der Lehrer benachteiligte einen Schüler konstant bei der Benotung.  
*Continuation:* Die Diskriminierung des Lehrers<sub>GenS</sub>/Schülers<sub>GenO</sub> zwang die Eltern zu einer Beschwerde.
70. *Context:* Der Beamte benachteiligte den älteren Bewerber bei der Vergabe des Postens.  
*Continuation:* Der Kollege bemerkte die Diskriminierung des Beamten<sub>GenS</sub>/Bewerbers<sub>GenO</sub> von Anfang an.

### **Durchleuchtung**

71. *Context:* Der Ermittler überprüfte den Verdächtigen wochenlang von Kopf bis Fuß.  
*Continuation:* Die Durchleuchtung des Ermittlers<sub>GenS</sub>/Verdächtigen<sub>GenO</sub> erbrachte den Beweis für eine Verurteilung.
72. *Context:* Der Staatsanwalt überprüfte den verdächtigen Politiker wieder und wieder.  
*Continuation:* Die Parteifreunde bezeichneten die Durchleuchtung des Staatsanwalts<sub>GenS</sub>/Politikers<sub>GenO</sub> als Hexenjagd.

### **Durchsuchung**

73. *Context:* Der Kommissar überprüfte den Räuber bei der Festnahme auf Waffen.  
*Continuation:* Die Durchsuchung des Kommissars<sub>GenS</sub>/Räubers<sub>GenO</sub> lockte einige Schaulustige in die Nähe.
74. *Context:* Der Zöllner kontrollierte bei der Einreise, ob der Fahrer Drogen bei sich hatte.  
*Continuation:* Ein Kollege beobachtete die Durchsuchung des Zöllners<sub>GenS</sub>/Fahrers<sub>GenO</sub> aus seinem Wagen heraus.

### **Einbürgerung**

75. *Context:* Der Minister sorgte dafür, dass der Sportler sofort die deutsche Staatsbürgerschaft erhielt.  
*Continuation:* Die Einbürgerung des Ministers<sub>GenS</sub>/Sportlers<sub>GenO</sub> erntete einigen Spott von der internationalen Presse.
76. *Context:* Der Richter sprach dem Sohn die Staatsbürgerschaft zu.  
*Continuation:* Die Familie feierte die Einbürgerung des Richters<sub>GenS</sub>/Sohnes<sub>GenO</sub> mit einem großen Fest.

### **Einlieferung**

77. *Context:* Der Notarzt brachte den Filmstar nach dem Zusammenbruch in die Klinik.  
*Continuation:* Die Einlieferung des Notarztes<sub>GenS</sub>/Filmstars<sub>GenO</sub> veranlasste die Presse zu wilden Spekulationen.
78. *Context:* Der Vater raste mit dem Mädchen nach ihrem Unfall in die Notaufnahme.  
*Continuation:* Ein Stau verzögerte die Einlieferung des Vaters<sub>GenS</sub>/Mädchens<sub>GenO</sub> um wenige Minuten.

### Einschüchterung

79. *Context:* Der Gangster drohte dem Zeugen vor dem Prozess.  
*Continuation:* Die Einschüchterung des Gangsters<sub>GenS</sub>/Zeugen<sub>GenO</sub> zeigte keine Wirkung während der Aussage.
80. *Context:* Der Minister drohte dem Reporter mit einer Anzeige wegen Verleumdung.  
*Continuation:* Die Zeitung kommentierte die Einschüchterung des Ministers<sub>GenS</sub>/Reporters<sub>GenO</sub> in einer Stellungnahme.

### Enteignung

81. *Context:* Der Richter beschlagnahmte das Vermögen des Betrügers für seine Opfer.  
*Continuation:* Die Enteignung des Richters<sub>GenS</sub>/Betrügers<sub>GenO</sub> gab den Opfern etwas Hoffnung.
82. *Context:* Das Ministerium verstaatlichte das Grundstück des Bauern für den Bau einer Straße.  
*Continuation:* Der Anwalt bekämpfte die Enteignung des Ministeriums<sub>GenS</sub>/Bauern<sub>GenO</sub> vor dem Höchstgericht.

### Enterbung

83. *Context:* Der Vater strich den Sohn aus dem Testament.  
*Continuation:* Die Enterbung des Vaters<sub>GenS</sub>/Sohnes<sub>GenO</sub> stürzte die Familie in einen Streit.
84. *Context:* Der Großvater beauftragte den Anwalt, den Enkel aus dem Testament zu nehmen.  
*Continuation:* Die Familie bekämpfte die Enterbung des Großvaters<sub>GenS</sub>/Enkels<sub>GenO</sub> vor Gericht.

### Entführung

85. *Context:* Der Terrorist hielt den Diplomat bis zur Lösegeldzahlung gefangen.  
*Continuation:* Die Entführung des Terroristen<sub>GenS</sub>/Diplomaten<sub>GenO</sub> hielt die Angehörigen in Atem.
86. *Context:* Der Täter verschleppte den Jungen in ein abgelegenes Versteck.  
*Continuation:* Die Familie bemerkte die Entführung des Täters<sub>GenS</sub>/Jungen<sub>GenO</sub> nach einigen Stunden.

### Enthauptung

87. *Context:* Der Entführer tötete den Gefangenen vor laufender Kamera.  
*Continuation:* Die Enthauptung des Entführers<sub>GenS</sub>/Gefangenen<sub>GenO</sub> schockierte die Ermittler beim Ansehen des Videos.
88. *Context:* Der Henker führte den Verurteilten zum Schafott.  
*Continuation:* Hunderte Menschen verfolgten die Enthauptung des Henkers<sub>GenS</sub>/Verurteilten<sub>GenO</sub> mit großem Entsetzen.

### Enthebung

89. *Context:* Der Vorstand der Partei entzog dem korrupten Abgeordneten alle Funktionen.  
*Continuation:* Die Enthebung des Vorstands<sub>GenS</sub>/Abgeordneten<sub>GenO</sub> fand viel Unterstützung in der Partei.
90. *Context:* Der Minister suspendierte den General, nachdem die Vorwürfe bekannt wurden.  
*Continuation:* Der Kanzler befürwortete die Enthebung des Ministers<sub>GenS</sub>/Generals<sub>GenO</sub> in einem Interview.

### Entlarvung

91. *Context:* Der Kommissar fand nach vielen Jahren den wahren Mörder des Mädchens.  
*Continuation:* Die Entlarvung des Kommissars<sub>GenS</sub>/Mörders<sub>GenO</sub> befreite den Vater von dem Mordverdacht.
92. *Context:* Der Reporter überführte den Radfahrer nach seinem Sieg des Dopings.  
*Continuation:* Die Polizei nutzte die Entlarvung des Reporters<sub>GenO</sub>/Radfahrers<sub>GenS</sub> für eine Razzia im Teambüro.

### Entlassung

93. *Context*: Der Chef kündigte dem Arbeiter nach zwölf Jahren in der Firma.  
*Continuation*: Die Entlassung des Chefs<sub>GenS</sub>/Arbeiters<sub>GenO</sub> brachte die Familie in eine schwierige Situation.
94. *Context*: Der Chirurg schickte den Jungen vier Monate nach der Operation nach Hause.  
*Continuation*: Die Entlassung des Chirurgen<sub>GenS</sub>/Jungen<sub>GenO</sub> freute das Personal auf der Station.

#### Entlastung

95. *Context*: Der Assistent nahm dem Chef viel Arbeit ab.  
*Continuation*: Die Entlastung des Assistenten<sub>GenS</sub>/Chefs<sub>GenO</sub> erleichterte die Arbeit an dem neuen Projekt.
96. *Context*: Der Praktikant half dem Angestellten drei Monate lang mit seiner Arbeit.  
*Continuation*: Der Chef begrüßte die Entlastung des Praktikanten<sub>GenS</sub>/Angestellten<sub>GenO</sub> sehr.

#### Entmachtung

97. *Context*: Der Diktator entzog dem Parlament nach dem Putsch alle Kompetenzen.  
*Continuation*: Die Entmachtung des Diktators<sub>GenS</sub>/Parlaments<sub>GenO</sub> bedeutete das Ende für die Demokratie.
98. *Context*: Das Volk stürzte das Regime mit einer friedlichen Revolution.  
*Continuation*: Die Armee unterstützte die Entmachtung des Volkes<sub>GenS</sub>/Regimes<sub>GenO</sub> durch Untätigkeit.

#### Entwaffnung

99. *Context*: Der mutige Kunde schlug dem Räuber die Pistole aus der Hand.  
*Continuation*: Die Entwaffnung des Kunden<sub>GenS</sub>/Räubers<sub>GenO</sub> zwang den Komplizen zu einer schnellen Flucht.
100. *Context*: Der Agent schlug dem Attentäter gerade noch rechtzeitig die Waffe aus der Hand.  
*Continuation*: Ein Reporter filmte die Entwaffnung des Agenten<sub>GenS</sub>/Attentäters<sub>GenO</sub> aus nächster Nähe.

#### Ergreifung

101. *Context*: Der Polizist schnappte den Mörder nach langen Ermittlungen.  
*Continuation*: Die Ergreifung des Polizisten<sub>GenS</sub>/Mörders<sub>GenO</sub> ermöglichte eine Verurteilung nach Jahren voller Rückschläge.
102. *Context*: Der Agent überwältigte den Terroristen nach einem kurzen Kampf.  
*Continuation*: Der Reporter schilderte die Ergreifung des Agenten<sub>GenS</sub>/Terroristen<sub>GenO</sub> in den Nachrichten.

#### Ermahnung

103. *Context*: Der Schiedsrichter wies den Tormann nach einem Foul zurecht.  
*Continuation*: Die Ermahnung des Schiedsrichters<sub>GenS</sub>/Tormanns<sub>GenO</sub> brachte die Fans zum Johlen.
104. *Context*: Der Trainer drohte dem Spieler, ihn bei weiteren Verfehlungen aus dem Kader zu nehmen.  
*Continuation*: Die Presse kommentierte die Ermahnung des Trainers<sub>GenS</sub>/Spielers<sub>GenO</sub> in zahllosen Artikeln.

#### Ermordung

105. *Context*: Der Attentäter tötete den Politiker während einer Wahlveranstaltung.  
*Continuation*: Die Ermordung des Attentäters<sub>GenS</sub>/Politikers<sub>GenO</sub> verzögerte die Wahl um einige Monate.
106. *Context*: Der Gangster erschoss den Richter, als er aus dem Haus kam.  
*Continuation*: Ein Zeuge schilderte die Ermordung des Richters<sub>GenS</sub>/Gangsters<sub>GenO</sub> mit zittriger Stimme.

#### Ermutigung

107. *Context*: Der Trainer richtete das Team nach der knappen Niederlage wieder etwas auf.  
*Continuation*: Die Ermutigung des Trainers<sub>GenS</sub>/Teams<sub>GenO</sub> machte die Niederlage für alle etwas erträglicher.
108. *Context*: Der Onkel bestärkte den Neffen in seinem Wunsch, zum Militär zu gehen.  
*Continuation*: Die Tante bezeichnete die Ermutigung des Onkels<sub>GenS</sub>/Neffen<sub>GenO</sub> als totalen Wahnsinn.

#### Erpressung

109. *Context:* Der ehemalige Assistent drohte dem Manager damit, belastende Daten weiterzuleiten.  
*Continuation:* Die Erpressung des Assistenten<sub>GenS</sub>/Managers<sub>GenO</sub> brachte die Konzernführung in eine schwierige Situation.
110. *Context:* Der Diener drohte dem Prinzen mit der Veröffentlichung von peinlichen Fotos.  
*Continuation:* Die Presse enthüllte die Erpressung des Dieners<sub>GenS</sub>/Prinzen<sub>GenO</sub> nach kurzer Zeit.

#### **Förderung**

111. *Context:* Der Millionär unterstützte den Künstler mit einer beträchtlichen Summe.  
*Continuation:* Die Förderung des Millionärs<sub>GenS</sub>/Künstlers<sub>GenO</sub> ermöglichte eine Arbeit ohne großen Druck.
112. *Context:* Der Konzern garantierte dem Verein aus der zweiten Liga Zahlungen über zwei Jahre.  
*Continuation:* Der Vorstand stoppte die Förderung des Konzerns<sub>GenS</sub>/Vereins<sub>GenO</sub> nach den zwei Jahren.

#### **Freilassung**

113. *Context:* Der Detektiv des Kaufhauses ließ den Jungen nach dem versuchten Diebstahl laufen.  
*Continuation:* Die Freilassung des Detektivs<sub>GenS</sub>/Jungen<sub>GenO</sub> ersparte der Familie eine Menge Ärger.
114. *Context:* Der Richter entließ den Entführer wegen Mangel an Beweisen aus der U-Haft.  
*Continuation:* Der Staatsanwalt kritisierte die Freilassung des Richters<sub>GenS</sub>/Entführers<sub>GenO</sub> in einer Stellungnahme.

#### **Heilung**

115. *Context:* Der Psychologe befreite den Patienten von seinen Angstzuständen.  
*Continuation:* Die Heilung des Psychologen<sub>GenS</sub>/Patienten<sub>GenO</sub> bedeutete einen Neubeginn für die ganze Familie.
116. *Context:* Der Arzt rettete den kranken Jungen mit einem neuen Medikament.  
*Continuation:* Das Medikament ermöglichte die Heilung des Arztes<sub>GenS</sub>/Jungen<sub>GenO</sub> ohne chirurgischen Eingriff.

#### **Inhaftierung**

117. *Context:* Der Inspektor steckte den Verdächtigen in Untersuchungshaft.  
*Continuation:* Die Inhaftierung des Inspektors<sub>GenS</sub>/Verdächtigen<sub>GenO</sub> brachte das Verfahren ins Rollen.
118. *Context:* Der Agent sollte den Terroristen in seiner Wohnung festnehmen.  
*Continuation:* Ein Fehler verhinderte die Inhaftierung des Agenten<sub>GenS</sub>/Terroristen<sub>GenO</sub> in letzter Minute.

#### **Manipulierung**

119. *Context:* Der Pate warnte den Zeugen vor einer Aussage gegen ihn.  
*Continuation:* Die Manipulierung des Paten<sub>GenS</sub>/Zeugen<sub>GenO</sub> gab dem Prozess eine unerwartete Wende.
120. *Context:* Der Gangster versprach dem Geschworenen im Fall eines Freispruchs eine hohe Summe.  
*Continuation:* Eine Reporterin enthüllte die Manipulierung des Gangsters<sub>GenS</sub>/Geschworenen<sub>GenO</sub> nach dem Ende des Prozesses.

#### **Rehabilitierung**

121. *Context:* Der Richter hob das Urteil gegen den Mann nach 12 Jahren auf.  
*Continuation:* Die Rehabilitierung des Richters<sub>GenS</sub>/Mannes<sub>GenO</sub> rührte den Anwalt zu Tränen.
122. *Context:* Der Detektiv bewies die Unschuld des Angeklagten vor Gericht.  
*Continuation:* Die Presse kommentierte die Rehabilitierung des Detektivs<sub>GenS</sub>/Angeklagten<sub>GenO</sub> in zahlreichen Berichten.

#### **Rettung**

123. *Context:* Der Passant holte den Fahrer in letzter Minute aus dem Autowrack.  
*Continuation:* Die Rettung des Passanten<sub>GenS</sub>/Fahrers<sub>GenO</sub> beherrschte die Titelblätter in den lokalen Zeitungen.

124. *Context:* Der Hund fand das Opfer der Lawine noch rechtzeitig unter dem Schnee.  
*Continuation:* Der Einsatzleiter beobachtete die Rettung des Hundes<sub>GenS</sub>/Opfers<sub>GenO</sub> aus dem Helikopter.

#### Schulung

125. *Context:* Der Experte erklärte dem Arbeiter in dem Seminar die neue Maschine.  
*Continuation:* Die Schulung des Experten<sub>GenS</sub>/Arbeiters<sub>GenO</sub> erleichterte die Arbeit an der Maschine um Einiges.
126. *Context:* Der Experte sollte das Team an einem Abend über Sicherheit am Arbeitsplatz aufklären.  
*Continuation:* Ein Vorfall verzögerte die Schulung des Experten<sub>GenS</sub>/Teams<sub>GenO</sub> um einige Tage.

#### Täuschung

127. *Context:* Der Verkäufer führte den Jungen beim Kauf hinters Licht.  
*Continuation:* Die Täuschung des Verkäufers<sub>GenS</sub>/Jungen<sub>GenO</sub> verärgerte den Vater ungemein.
128. *Context:* Der Betrüger legte den Gutachter der Versicherung herein.  
*Continuation:* Ein Detektiv enthüllte die Täuschung des Betrügers<sub>GenS</sub>/Gutachters<sub>GenO</sub> nach einiger Zeit.

#### Tötung

129. *Context:* Der Agent erschoss den Attentäter nach dem mißglückten Anschlag.  
*Continuation:* Die Tötung des Agenten<sub>GenS</sub>/Attentäters<sub>GenO</sub> verursachte große Panik unter den Anwesenden.
130. *Context:* Der Wilderer erschoss den Löwen mitten im Nationalpark.  
*Continuation:* Ein Ranger beobachtete die Tötung des Wilderers<sub>GenS</sub>/Löwen<sub>GenO</sub> aus der Ferne.

#### Überprüfung

131. *Context:* Der Beamte prüfte den Hintergrund des Touristen vor der Einreise per Computer.  
*Continuation:* Die Überprüfung des Beamten<sub>GenS</sub>/Touristen<sub>GenO</sub> verzögerte die Einreise um einige Minuten.
132. *Context:* Der Detektiv durchleuchtete den Verdächtigen von Kopf bis Fuß.  
*Continuation:* Der Auftraggeber bezahlte die Überprüfung des Detektivs<sub>GenS</sub>/Verdächtigen<sub>GenO</sub> im Voraus.

#### Überstellung

133. *Context:* Der Wärter begleitete den Häftling in ein neues Gefängnis.  
*Continuation:* Die Überstellung des Wärters<sub>GenS</sub>/Häftlings<sub>GenO</sub> ermöglichte eine Verwahrung ohne Sicherheitsrisiko.
134. *Context:* Der Pfleger brachte den Patienten von der Intensivstation in ein normales Zimmer.  
*Continuation:* Der Arzt bestätigte die Überstellung des Pflegers<sub>GenS</sub>/Patienten<sub>GenO</sub> mit seiner Unterschrift.

#### Überwachung

135. *Context:* Der Agent folgte dem Diplomaten über mehrere Wochen.  
*Continuation:* Die Überwachung des Agenten<sub>GenS</sub>/Diplomaten<sub>GenO</sub> erbrachte keine Hinweise auf mögliche Spionage.
136. *Context:* Der Detektiv beobachtete den Verdächtigen über mehrere Monate.  
*Continuation:* Der Auftraggeber bezahlte die Überwachung des Detektivs<sub>GenS</sub>/Verdächtigen<sub>GenO</sub> mit seinen Ersparnissen.

#### Umarmung

137. *Context:* Der Vater schloss den Sohn nach 20 Jahren Streit in seine Arme.  
*Continuation:* Die Umarmung des Vaters<sub>GenS</sub>/Sohnes<sub>GenS</sub> rührte die Mutter zu Tränen.
138. *Context:* Der Papst schloss den Pilger in seine Arme.  
*Continuation:* Der Regisseur zeigte die Umarmung des Papstes<sub>GenS</sub>/Pilgers<sub>GenO</sub> in einer Nahaufnahme.

#### Unterdrückung

139. *Context:* Der Tyrann bekämpfte den Widerstand mit brutalen Mitteln.  
*Continuation:* Die Unterdrückung des Tyrannen<sub>GenS</sub>/Widerstandes<sub>GenO</sub> forderte viele Opfer unter den Rebellen.
140. *Context:* Der Diktator beutete das Volk brutal und gnadenlos aus.  
*Continuation:* Ein Putsch beendete die Unterdrückung des Diktators<sub>GenS</sub>/Volkes<sub>GenO</sub> nach drei Jahren.

#### Unterstützung

141. *Context:* Der Konzern förderte den Sportler mit einer großen Summe.  
*Continuation:* Die Unterstützung des Konzerns<sub>GenS</sub>/Sportlers<sub>GenO</sub> hatte positive Auswirkungen auf das Image der Firma.
142. *Context:* Der Senat bewilligte dem umstrittenen Künstler einen großen Zuschuss für seine Aktion.  
*Continuation:* Eine Zeitung kritisierte die Unterstützung des Senats<sub>GenS</sub>/Künstlers<sub>GenO</sub> auf dem Titelblatt.

#### Untersuchung

143. *Context:* Der Doktor überprüfte den Mann auf mögliche neurologische Störungen.  
*Continuation:* Die Untersuchung des Doktors<sub>GenS</sub>/Mannes<sub>GenO</sub> brachte keinen Hinweis auf eine Erkrankung.
144. *Context:* Der Ermittler der Steuerbehörde überprüfte den Unternehmer nach einem Hinweis.  
*Continuation:* Der Vorgesetzte genehmigte die Untersuchung des Ermittlers<sub>GenS</sub>/Unternehmers<sub>GenO</sub> mit seiner Unterschrift.

#### Verbannung

145. *Context:* Der König schickte den aufständischen General ins Exil.  
*Continuation:* Die Verbannung des Königs<sub>GenS</sub>/Generals<sub>GenO</sub> beendete den Aufstand gegen die Monarchie.
146. *Context:* Der Bischof schickte den ketzerischen Mönch ins Exil.  
*Continuation:* Eine Urkunde belegt die Verbannung des Bischofs<sub>GenS</sub>/Mönchs<sub>GenO</sub> in allen Details.

#### Verhaftung

147. *Context:* Der Agent fasste den Terroristen nach einem Tipp aus der Bevölkerung.  
*Continuation:* Die Verhaftung des Agenten<sub>GenS</sub>/Terroristen<sub>GenO</sub> verhinderte einen Anschlag auf einen Zug.
148. *Context:* Der Polizist schnappte den Mörder nach einer wilden Verfolgungsjagd.  
*Continuation:* Der Reporter filmte die Verhaftung des Polizisten<sub>GenS</sub>/Mörders<sub>GenO</sub> für die Abendnachrichten.

#### Verlegung

149. *Context:* Der Pfleger brachte den Patienten von der Intensivstation in die Chirurgie.  
*Continuation:* Die Verlegung des Pflegers<sub>GenS</sub>/Patienten<sub>GenO</sub> überraschte den Chefarzt der Chirurgie.
150. *Context:* Der Wächter wollte den Gefangenen in einen anderen Trakt bringen.  
*Continuation:* Der Direktor bewilligte die Verlegung des Wächters<sub>GenS</sub>/Gefangenen<sub>GenO</sub> nach einem Gespräch.

#### Vernehmung

151. *Context:* Der Kommissar befragte den verhafteten Kriminellen über die Organisation.  
*Continuation:* Die Vernehmung des Kommissars<sub>GenS</sub>/Kriminellen<sub>GenO</sub> brachte viele Informationen für den Staatsanwalt.
152. *Context:* Der Staatsanwalt befragte den Kronzeugen im Zeugenstand.  
*Continuation:* Der Verteidiger störte die Vernehmung des Staatsanwalts<sub>GenS</sub>/Kronzeugen<sub>GenO</sub> mit vielen Einsprüchen.

#### Verschleppung



153. *Context:* Der Spion entführte den Forscher aus seinem Heimatland.  
*Continuation:* Die Verschleppung des Spions<sub>GenS</sub>/Forschers<sub>GenO</sub> verschärfte die Spannungen zwischen den Ländern.
154. *Context:* Der Vater brachte das Kind ohne Wissen der Mutter in sein Heimatland.  
*Continuation:* Die Mutter meldete die Verschleppung des Vaters<sub>GenS</sub>/Kindes<sub>GenO</sub> an die Polizei.

#### **Versetzung**

155. *Context:* Der General wies den Soldaten einer neuen Einheit zu.  
*Continuation:* Die Versetzung des Generals<sub>GenS</sub>/Soldaten<sub>GenO</sub> überraschte die Kollegen in der Kompanie.
156. *Context:* Der Chef schickte den Mitarbeiter zum Aufbau des neuen Büros nach Asien.  
*Continuation:* Die Kollegen bedauerten die Versetzung des Chefs<sub>GenS</sub>/Mitarbeiters<sub>GenO</sub> nach Asien.

#### **Versorgung**

157. *Context:* Der Sponsor sorgte während des zweiwöchigen Radrennens dafür, dass der Fahrer gut betreut wurde.  
*Continuation:* Die Versorgung des Sponsors<sub>GenS</sub>/Fahrers<sub>GenO</sub> erleichterte die Arbeit für das gesamte Team.
158. *Context:* Der Koch stellte während der Expedition sicher, dass das Team gut verpflegt wurde.  
*Continuation:* Der Leiter lobte die Versorgung des Kochs<sub>GenS</sub>/Teams<sub>GenO</sub> nach dem Abenteuer.

#### **Verteidigung**

159. *Context:* Der Anwalt vertrat den Manager in einem schwierigen Prozess.  
*Continuation:* Die Verteidigung des Anwalts<sub>GenS</sub>/Managers<sub>GenO</sub> erforderte eine Strategie mit viel Risiko.
160. *Context:* Das Team hielt den Stürmer das ganze Spiel über unter Kontrolle.  
*Continuation:* Der Trainer lobte die Verteidigung des Teams<sub>GenS</sub>/Stürmers<sub>GenO</sub> nach dem Spiel.

#### **Vertreibung**

161. *Context:* Der Besitzer zwang den Mieter, aus der Wohnung auszuziehen.  
*Continuation:* Die Vertreibung des Besitzers<sub>GenS</sub>/Mieters<sub>GenO</sub> veranlasste den Anwalt zu einer Klage.
162. *Context:* Der General jagte den König nach dem Putsch aus dem Land.  
*Continuation:* Die UNO verurteilte die Vertreibung des Generals<sub>GenS</sub>/Königs<sub>GenO</sub> in einer Stellungnahme.

#### **Verurteilung**

163. *Context:* Der Richter schickte den Mann für zehn Jahre ins Gefängnis.  
*Continuation:* Die Verurteilung des Richters<sub>GenS</sub>/Mannes<sub>GenO</sub> überraschte die Zuseher im Gerichtssaal.
164. *Context:* Das Gremium sperrte den Sportler nach dem positiven Dopingtests für zwei Jahre.  
*Continuation:* Der Verband verkündete die Verurteilung des Gremiums<sub>GenS</sub>/Sportlers<sub>GenO</sub> in einer Pressemitteilung.

#### **Würdigung**

165. *Context:* Der Vorgesetzte lobte den Angestellten in der Besprechung in den Himmel.  
*Continuation:* Die Würdigung des Vorgesetzten<sub>GenS</sub>/Angestellten<sub>GenO</sub> beeindruckte die Kollegen in der Firma.
166. *Context:* Der Minister betonte in seiner Rede die Verdienste des Künstlers.  
*Continuation:* Das Fernsehen zeigte die Würdigung des Ministers<sub>GenS</sub>/Künstlers<sub>GenO</sub> in einem Beitrag.

#### **Zurückweisung**

167. *Context:* Der Großvater akzeptierte seinen unehelichen Enkel nie.  
*Continuation:* Die Zurückweisung des Großvaters<sub>GenS</sub>/Enkels<sub>GenO</sub> belastete die Familie über viele Jahre hinweg.
168. *Context:* Der Stiefvater ließ den Jungen nie an sich heran.  
*Continuation:* Die Mutter akzeptierte die Zurückweisung des Stiefvaters<sub>GenS</sub>/Jungen<sub>GenO</sub> schweigend.



## Appendix D

# Supplementary materials for the event-related potentials study

### Supplementary analyses

#### Rating data

The following tables summarise the additional cumulative link mixed models (CLMMs) fitted on the rating data of the SPR experiment using function `clmm` from the *ordinal* package (Christensen, 2013b). Table D.1 presents the estimates for the full data set, while tables D.2 and D.3 summarise the CLMMs fitted on the GenO and GenS subsets, respectively.

#### Random effects

Groups	Name	Variance	SD	Corr.
Story	(Intercept)	0.14	0.37	
	c.Linking	1.43	1.20	0.17
Subject	(Intercept)	0.20	0.44	
	c.Linking	1.31	1.14	-0.25
	c.AI	0.01	0.09	0.62 0.46

#### Fixed effects

	Estimate	SE	Lower	Upper	z-value	Pr(> z )
c.Linking	-2.88	0.25	-3.38	-2.39	-11.42	<0.001 ***
c.AI	-0.14	0.03	-0.20	-0.07	-3.89	<0.001 ***
c.Linking:c.AI	-1.19	0.08	-1.35	-1.04	-14.96	<0.001 ***

**Table D.1:** Summary tables for the cumulative link mixed model (CLMM) for the acceptability rating data of the ERP experiment.

Random effects

Groups	Name	Variance	SD	Corr.
Story	(Intercept)	0.43	0.65	
Subject	(Intercept)	0.66	0.81	
	c.AI	0.03	0.17	0.50

Fixed effects

	Estimate	SE	Lower	Upper	z-value	Pr(> z )	
c.AI	0.47	0.06	0.36	0.59	8.11	<0.001	***

**Table D.2:** CLMM results for the GenO subset of the ERP acceptability data.

Random effects

Groups	Name	Variance	SD	Corr.
Story	(Intercept)	0.61	0.78	
Subject	(Intercept)	0.43	0.66	
	c.AI	0.04	0.21	0.91

Fixed effects

	Estimate	SE	Lower	Upper	z-value	Pr(> z )	
c.AI	-0.75	0.07	-0.88	-0.62	-11.05	<0.001	***

**Table D.3:** CLMM results for the GenS subset of the ERP acceptability data.

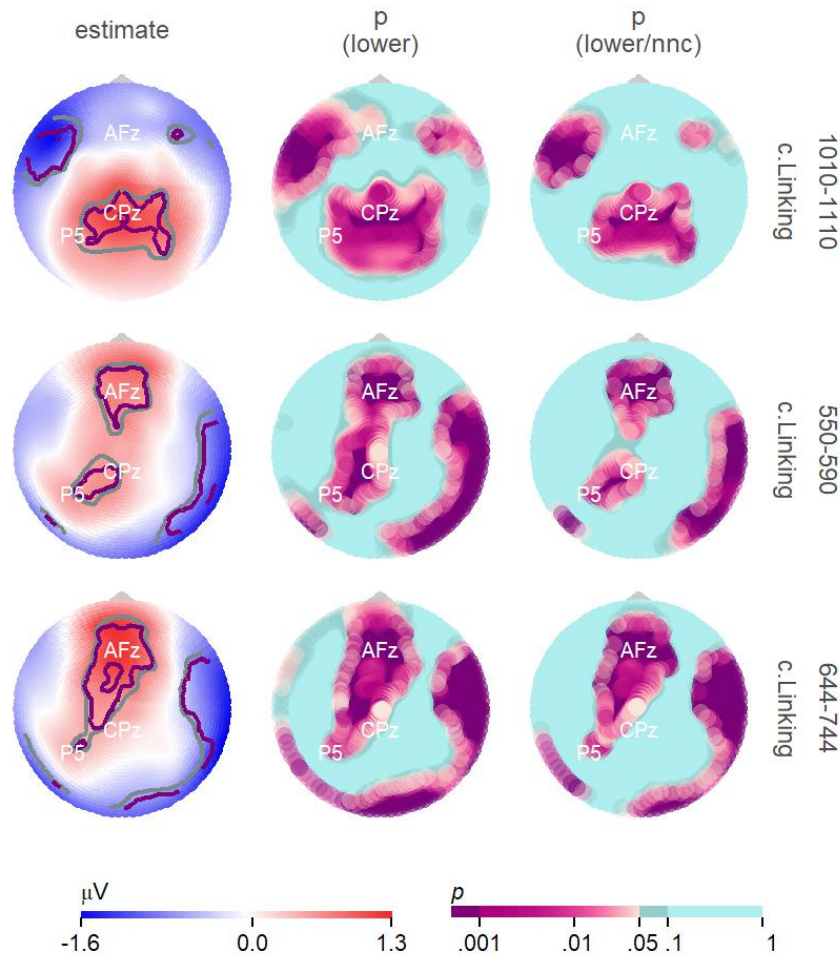
## ERP data

The following plots present the supplementary ERP-analyses for the most relevant findings based upon the `lmerMom` function of the *stepmom* package, focussing on the genitive segment only. For these, time windows were predefined based upon the periods of peak effects identified using the `wfmm`-based analysis (see table 4.14) and the resulting mean-ERP values were used as dependent variables (for more details see section 4.2.6). The model run separately for each electrode and time-window mean ERP was defined as in (130) below:

(130) *Genitive segment:*  
Mean-ERP ~ c.Linking \* s.AI + c.Embedding + s.List.Pos +  
s.Word.Length + s.log.Frequency +  
(1 | Subject) + (0 + c.Linking | Subject) +  
(0 + s.AI | Subject) + (0 + c.Linking:s.AI | Subject) + (1 | Word)

Figure D.1 presents the model maps of the 'c.Linking' predictor, corresponding to the time windows of figure 4.10; the leftmost column shows the maps for the estimates, the middle column the uncorrected lower-bound degrees of freedom *p*-values and the rightmost column the nearest-neighbours corrected *p*-values (see section 4.2.6). The middle row shows the

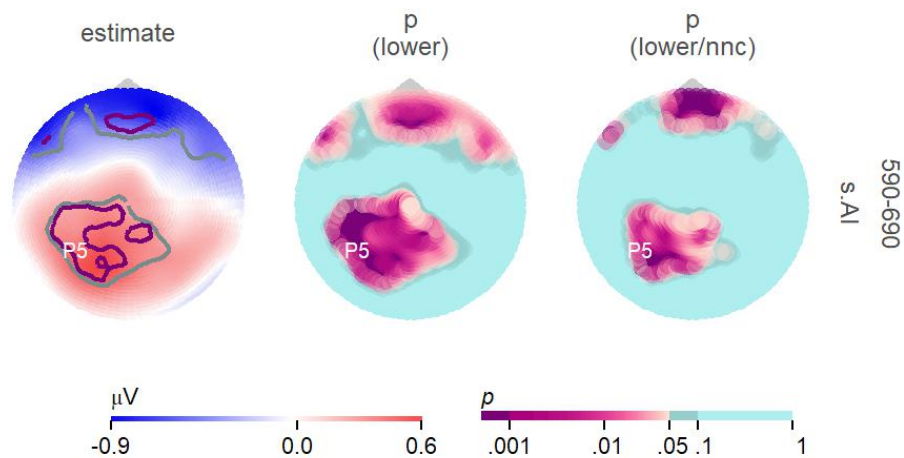
estimates for the time window from 550–590 ms post onset, illustrating the left posterior positivity associated with 'c.Linking' (corresponding to 4.10a); the bottom row shows the same for the time period from 644–744 ms, focussing on the peak of the anterior midline positivity (corresponding to 4.10b); the top row shows the model maps for the posterior midline positivity in the trailing time window from 1010–1110 ms (corresponding to 4.10c).



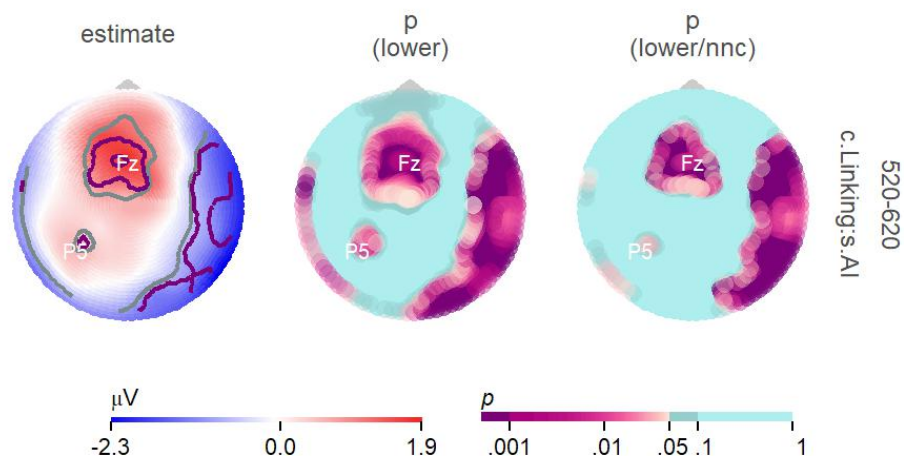
**Figure D.1:** Genitive segment - lmer-based Model maps for the 'c.Linking' predictor. The maps can be compared to the wfm-based maps presented in 4.10. For more details, please refer to the text.

Figure D.2 presents the model maps of the 's.AI' predictor: it illustrates the left posterior positivity in the time window from 590–690 ms, corresponding to the time windows of figure 4.11.

Figure D.3 presents the model maps of the 'c.Linking:s.AI' interaction predictor, illustrating the anterior midline positivity and the left posterior positivity in the time window from 520–620 ms, corresponding to the time windows of figure 4.12a.



**Figure D.2:** Genitive segment - 1mer-based Model maps for the 'a.AI' predictor. The maps can be compared to the wfmm-based maps presented in 4.11.



**Figure D.3:** Genitive segment - 1mer-based model maps for the 'c.Linking:s.AI' interaction predictor. The maps can be compared to the wfmm-based maps presented in 4.12 a.

## Stimulus sentences

The following pages list the stimulus sentences used in the experiment, alphabetically ordered by the relevant *-ung*-nominalisation and numbered consecutively by 'story'. The first sentence represents the context sentence, the second the continuation sentence; GenO and GenS subscripts indicate the classifications of the genitives presented in the continuation sentences.

### Ablehnung

1. *Context*: Der Direktor weigerte sich, den Forscher einzustellen.

*Continuation:* Die Ablehnung des Direktors<sub>GenS</sub>/Forschers<sub>GenO</sub> enttäuschte die Kollegen im Labor.

2. *Context:* Der Papst verweigerte dem Kardinal das hohe Amt im Vatikan.

*Continuation:* Viele Journalisten bezeichneten die Ablehnung des Papstes<sub>GenS</sub>/Kardinals<sub>GenO</sub> als falsches Signal.

#### **Abschirmung**

3. *Context:* Der Anwalt hielt die Presse von dem Opfer fern.

*Continuation:* Die Abschirmung des Anwalts<sub>GenS</sub>/Opfers<sub>GenO</sub> ermöglichte die Vorbereitung auf den Prozess.

4. *Context:* Der Bodyguard ließ während des gesamten Termins niemanden an den Sportler heran.

*Continuation:* Die Fans hielten die Abschirmung des Bodyguards<sub>GenS</sub>/Sportlers<sub>GenO</sub> für übertrieben.

#### **Absetzung**

5. *Context:* Der Kanzler entließ seinen Minister, nachdem der Skandal aufgefliegen war.

*Continuation:* Die Absetzung des Kanzlers<sub>GenS</sub>/Ministers<sub>GenO</sub> rettete die Regierung vor einer Krise.

6. *Context:* Der Rektor enthob den Professor nach dem Skandal von seinem Posten.

*Continuation:* Die Studenten unterstützten die Absetzung des Rektors<sub>GenS</sub>/Professors<sub>GenO</sub> zum größten Teil.

#### **Abwerbung**

7. *Context:* Der Chef des Konzerns brachte den Forscher dazu, in sein Unternehmen zu wechseln.

*Continuation:* Die Abwerbung des Chefs<sub>GenS</sub>/Forschers<sub>GenO</sub> hatte ein Nachspiel vor Gericht.

8. *Context:* Der Manager des Fußballclubs kaufte den Star des Erzrivalen für seinen eigenen Verein.

*Continuation:* Die Fans quittierten die Abwerbung des Managers<sub>GenS</sub>/Stars<sub>GenO</sub> im nächsten Spiel mit Pfeifkonzerten.

#### **Anhörung**

9. *Context:* Der Vorstand befragte den Manager, um seine Sicht der Dinge zu hören.

*Continuation:* Die Anhörung des Vorstands<sub>GenS</sub>/Managers<sub>GenO</sub> erbrachte keine Hinweise auf einen Betrug.

10. *Context:* Der Senat befragte den Minister zu den Vorgängen.

*Continuation:* Das Fernsehen übertrug die Anhörung des Senats<sub>GenS</sub>/Ministers<sub>GenO</sub> in voller Länge.

#### **Aufforderung**

11. *Context:* Der liberale Bischof rief den Papst dazu auf, das Zölibat abzuschaffen.

*Continuation:* Die Aufforderung des Bischofs<sub>GenS</sub>/Papstes<sub>GenO</sub> veranlasste einen Reporter zu einem Interview.

12. *Context:* Der Politiker appelierte an den Manager, sich an den Kosten der Krise zu beteiligen.

*Continuation:* Die Presse kommentierte die Aufforderung des Politikers<sub>GenS</sub>/Managers<sub>GenO</sub> ziemlich pessimistisch.

#### **Aufmunterung**

13. *Context:* Der Vater tröstete das Kind nach der verpatzten Klausur.

*Continuation:* Die Aufmunterung des Vaters<sub>GenS</sub>/Kindes<sub>GenO</sub> rettete das Wochenende für die ganze Familie.

14. *Context:* Der Clown brachte den kleinen Patienten nach langer Zeit wieder zum Lachen.

*Continuation:* Die Ärztin verfolgte die Aufmunterung des Clowns<sub>GenS</sub>/Patienten<sub>GenO</sub> aus dem Hintergrund.

#### **Ausbeutung**

15. *Context:* Der Professor übertrug dem Assistenten jahrelang seine eigene Arbeit.

*Continuation:* Die Ausbeutung des Professors<sub>GenS</sub>/Assistenten<sub>GenO</sub> verärgerte viele Kollegen am Institut.

16. *Context:* Der Wirt ließ seinen Kellner für zu wenig Geld viel zu viel arbeiten.

*Continuation:* Die Kündigung beendete die Ausbeutung des Wirts<sub>GenS</sub>/Kellners<sub>GenO</sub> nach fast drei Jahren.

### Ausbildung

17. *Context:* Der Koch brachte dem Azubi im Lauf der Zeit viele neue Sachen bei.  
*Continuation:* Die Ausbildung des Kochs<sub>GenS</sub>/Azubis<sub>GenO</sub> beeindruckte den Chef des kleinen Restaurants.
18. *Context:* Der Meister vermittelte dem Lehrling sein ganzes Können mit viel Begeisterung.  
*Continuation:* Die Eltern lobten die Ausbildung des Meisters<sub>GenS</sub>/Lehrlings<sub>GenO</sub> bei jeder Gelegenheit.

### Auslieferung

19. *Context:* Der Senat schickte den geflohenen Diktator zurück in sein Land.  
*Continuation:* Die Auslieferung des Senats<sub>GenS</sub>/Diktators<sub>GenO</sub> verbesserte die Stimmung zwischen den Staaten.
20. *Context:* Der Richter überstellte den Verdächtigen nach langem Zögern in seine Heimat.  
*Continuation:* Die Behörden begrüßten die Auslieferung des Richters<sub>GenS</sub>/Verdächtigen<sub>GenO</sub> sehr.

### Auswechslung

21. *Context:* Der Kläger suchte sich kurz vor dem Prozess einen neuen Anwalt.  
*Continuation:* Die Auswechslung des Klägers<sub>GenS</sub>/Anwalts<sub>GenO</sub> verzögerte den Prozessbeginn um einige Wochen.
22. *Context:* Der Trainer nahm den Stürmer kurz nach der Halbzeitpause vom Feld.  
*Continuation:* Die Fans begleiteten die Auswechslung des Trainers<sub>GenS</sub>/Stürmers<sub>GenO</sub> mit einem Pfeifkonzert.

### Ausweisung

23. *Context:* Der Beamte lehnte den Antrag des Flüchtlings auf Asyl ab.  
*Continuation:* Die Ausweisung des Beamten<sub>GenS</sub>/Flüchtlings<sub>GenO</sub> schockierte viele Leute wegen ihrer Härte.
24. *Context:* Der Richter entschied, den Asylanten in sein Heimatland zurückzuschicken.  
*Continuation:* Der Anwalt bekämpfte die Ausweisung des Richters<sub>GenS</sub>/Asylanten<sub>GenO</sub> in der Revision.

### Beförderung

25. *Context:* Der Manager ernannte den Assistenten zum Projektleiter.  
*Continuation:* Die Beförderung des Managers<sub>GenS</sub>/Assistenten<sub>GenO</sub> freute die Kollegen im Büro.
26. *Context:* Der Minister ernannte den Beamten zum Botschafter.  
*Continuation:* Die Kollegen kritisierten die Beförderung des Ministers<sub>GenS</sub>/Beamten<sub>GenO</sub> intern.

### Befragung

27. *Context:* Der Polizist verhörte den Zeugen zum Ablauf des Überfalls.  
*Continuation:* Die Befragung des Polizisten<sub>GenS</sub>/Zeugen<sub>GenO</sub> erbrachte einen Hinweis auf das Fluchtauto.
28. *Context:* Der Staatsanwalt nahm den Angeklagten ins Kreuzverhör.  
*Continuation:* Der Verteidiger unterbrach die Befragung des Staatsanwalts<sub>GenS</sub>/Angeklagten<sub>GenO</sub> mit mehreren Einsprüchen.

### Befreiung

29. *Context:* Der Kommissar holte den entführten Jungen aus seinem Verlies.  
*Continuation:* Die Befreiung des Kommissars<sub>GenS</sub>/Jungen<sub>GenO</sub> brachte den Fall in die internationalen Medien.
30. *Context:* Das Team der Spezialeinheit erlöste den Politiker aus der Geiselhaft.  
*Continuation:* Ein Offizier schilderte die Befreiung des Teams<sub>GenS</sub>/Politikers<sub>GenO</sub> in einem Interview.

### Begrüßung

31. *Context:* Der Sänger richtete zu Beginn einige nette Worte an das Publikum.  
*Continuation:* Die Begrüßung des Sängers<sub>GenS</sub>/Publikums<sub>GenO</sub> eröffnete das Konzert in dem kleinen Club.



32. *Context:* Der Minister hieß den Athleten nach seinem Olympiasieg persönlich willkommen.  
*Continuation:* Der Reporter kommentierte die Begrüßung des Ministers<sub>GenS</sub>/Athleten<sub>GenO</sub> im Radio.

#### **Behandlung**

33. *Context:* Der Therapeut massierte den Sportler mit einer neuen Technik.  
*Continuation:* Die Behandlung des Therapeuten<sub>GenS</sub>/Sportlers<sub>GenO</sub> verringerte die Schmerzen nach wenigen Sitzungen.
34. *Context:* Der Hausarzt therapierte den Kranken mit einer alternativen Methode.  
*Continuation:* Die Kasse bezahlte die Behandlung des Hausarztes<sub>GenS</sub>/Kranken<sub>GenO</sub> zur Hälfte.

#### **Bekehrung**

35. *Context:* Der Priester überzeugte den Häftling davon, in die Kirche einzutreten.  
*Continuation:* Die Bekehrung des Priesters<sub>GenS</sub>/Häftlings<sub>GenO</sub> brachte den Direktor zum Staunen.
36. *Context:* Der Therapeut überzeugte den Raucher davon, mit dem Qualmen aufzuhören.  
*Continuation:* Die Familie bezeichnete die Bekehrung des Therapeuten<sub>GenS</sub>/Rachters<sub>GenO</sub> als ein kleines Wunder.

#### **Belagerung**

37. *Context:* Der Fan campierte monatelang vor dem Grundstück des Sängers.  
*Continuation:* Die Belagerung des Fans<sub>GenS</sub>/Sängers<sub>GenO</sub> zwang die Polizei nach einiger Zeit zum Handeln.
38. *Context:* Der Kaiser schnitt mit seiner Armee den Papst lange komplett von der Außenwelt ab.  
*Continuation:* Ein Kompromiss beendete die Belagerung des Kaisers<sub>GenS</sub>/Papstes<sub>GenO</sub> nach sechs Monaten.

#### **Belästigung**

39. *Context:* Ein Betrunkener provozierte den Wächter in der U-Bahn-Station.  
*Continuation:* Die Belästigung des Betrunkenen<sub>GenS</sub>/Wächters<sub>GenO</sub> veranlasste einen Fahrgast zum Eingreifen.
40. *Context:* Der Jugendliche bedrängte den Pensionisten auf offener Straße.  
*Continuation:* Ein Passant bemerkte die Belästigung des Jugendlichen<sub>GenS</sub>/Pensionisten<sub>GenO</sub> gerade noch rechtzeitig.

#### **Belehrung**

41. *Context:* Der Chef wies den nachlässigen Arbeiter auf seine Pflichten hin.  
*Continuation:* Die Belehrung des Chefs<sub>GenS</sub>/Arbeiters<sub>GenO</sub> verbesserte die Situation in den folgenden Wochen.
42. *Context:* Der Anwalt klärte den Angeklagten vor dem Prozess über seine Rechte auf.  
*Continuation:* Der Assistent unterbrach die Belehrung des Anwalts<sub>GenS</sub>/Angeklagten<sub>GenO</sub> wegen eines dringenden Telefonats.

#### **Beleidigung**

43. *Context:* Der Schüler beschimpfte den Lehrer in der Pause.  
*Continuation:* Die Beleidigung des Schülers<sub>GenS</sub>/Lehrers<sub>GenO</sub> schockierte die Mitschüler aus seiner Klasse.
44. *Context:* Der Fahrer zeigte dem Polizisten während einer Kontrolle den Mittelfinger.  
*Continuation:* Ein Kollege bestätigte die Beleidigung des Fahrers<sub>GenS</sub>/Polizisten<sub>GenO</sub> vor Gericht.

#### **Beratung**

45. *Context:* Der Experte half dem Vorstand des Unternehmens mit dem neuen Projekt.  
*Continuation:* Die Beratung des Experten<sub>GenS</sub>/Vorstands<sub>GenO</sub> kostete das Unternehmen ein kleines Vermögen.

46. *Context*: Der Fachmann erstellte dem Kandidaten vor dem Wahlkampf eine umfassende Analyse.  
*Continuation*: Die Partei bezahlte die Beratung des Fachmanns<sub>GenS</sub>/Kandidaten<sub>GenO</sub> aus der Parteikasse.

#### **Bergung**

47. *Context*: Der Sanitäter zog den Fahrer vorsichtig aus dem Autowrack.  
*Continuation*: Die Bergung des Sanitäters<sub>GenS</sub>/Fahrsers<sub>GenO</sub> ermöglichte dem Notarzt eine schnelle Behandlung.
48. *Context*: Der Pilot des Hubschraubers zog den gekenterten Segler aus der stürmischen See.  
*Continuation*: Die Küstenwache meldete die Bergung des Piloten<sub>GenS</sub>/Seglers<sub>GenO</sub> an die Angehörigen.

#### **Berührung**

49. *Context*: Der Tierarzt fasste dem kranken Pferd zum Kennenlernen vorsichtig an die Flanke.  
*Continuation*: Die Berührung des Tierarztes<sub>GenS</sub>/Pferdes<sub>GenO</sub> schaffte genug Vertrauen für die Behandlung.
50. *Context*: Der Vater berührte das Baby kurz nach der Geburt.  
*Continuation*: Der Großvater filmte die Berührung des Vaters<sub>GenS</sub>/Babys<sub>GenO</sub> mit seiner Kamera.

#### **Beschattung**

51. *Context*: Der Agent beobachtete den Politiker über mehrere Wochen.  
*Continuation*: Die Beschattung des Agenten<sub>GenS</sub>/Politikers<sub>GenO</sub> brachte viele Informationen für die geplante Aktion.
52. *Context*: Der Ermittler ließ den Dealer nicht aus den Augen.  
*Continuation*: Der General koordinierte die Beschattung des Ermittlers<sub>GenS</sub>/Dealers<sub>GenO</sub> persönlich.

#### **Beschimpfung**

53. *Context*: Der Chef verletzte den Arbeiter mit seinen Beleidigungen.  
*Continuation*: Die Beschimpfung des Chefs<sub>GenS</sub>/Arbeiters<sub>GenO</sub> verschlechterte die Stimmung noch weiter.
54. *Context*: Der Stürmer beleidigte den Tormann nach dem gehaltenen Elfmeter.  
*Continuation*: Der Schiedsrichter ahndete die Beschimpfung des Stürmers<sub>GenS</sub>/Tormanns<sub>GenO</sub> mit der roten Karte.

#### **Beseitigung**

55. *Context*: Der Agent vergiftete den Diktator beim Abendessen.  
*Continuation*: Die Beseitigung des Agenten<sub>GenS</sub>/Diktators<sub>GenO</sub> öffnete den Weg zu einem Machtwechsel.
56. *Context*: Der Spekulant trieb den Betrieb gezielt in den Konkurs.  
*Continuation*: Ein Journalist enthüllte die Beseitigung des Spekulanten<sub>GenS</sub>/Betriebs<sub>GenO</sub> in einem Artikel.

#### **Bestechung**

57. *Context*: Der Manager bot dem zuständigen Beamten eine Menge Geld für den Bauauftrag.  
*Continuation*: Die Bestechung des Managers<sub>GenS</sub>/Beamten<sub>GenO</sub> entschied den Wettbewerb zu Gunsten der Firma.
58. *Context*: Der Vorstand bezahlte den Betriebsrat für sein Schweigen.  
*Continuation*: Ein Kollege meldete die Bestechung des Vorstands<sub>GenS</sub>/Betriebsrats<sub>GenO</sub> an den Staatsanwalt.

#### **Bestrafung**

59. *Context*: Der Erzieher stellte den Jungen für eine halbe Stunde in die Ecke.  
*Continuation*: Die Bestrafung des Erziehers<sub>GenS</sub>/Jungen<sub>GenO</sub> veranlasste die Eltern zu einer Beschwerde.
60. *Context*: Der Richter verurteilte den Täter nur zu mehreren Wochen Sozialarbeit.  
*Continuation*: Die Presse hielt die Bestrafung des Richters<sub>GenS</sub>/Täters<sub>GenO</sub> für viel zu milde.

#### **Betreuung**

61. *Context*: Der Pfleger beaufsichtigte den behinderten Jungen tagsüber.  
*Continuation*: Die Betreuung des Pflegers<sub>GenS</sub>/Jungen<sub>GenO</sub> erleichterte das Leben für die Eltern ungemein.
62. *Context*: Der Dolmetscher kümmerte sich während der gesamten Geschäftsreise um den Manager.  
*Continuation*: Der Gastgeber bezahlte die Betreuung des Dolmetschers<sub>GenS</sub>/Managers<sub>GenO</sub> aus der eigenen Tasche.

#### **Bevormundung**

63. *Context*: Der Arzt verweigerte dem Patienten die übliche Behandlung.  
*Continuation*: Die Bevormundung des Arztes<sub>GenS</sub>/Patienten<sub>GenO</sub> hatte ein Nachspiel vor Gericht.
64. *Context*: Der Vater entschied, welche Ausbildung der Sohn machen sollte.  
*Continuation*: Der Großvater kritisierte die Bevormundung des Vaters<sub>GenS</sub>/Sohnes<sub>GenO</sub> bei einem gemeinsamen Essen.

#### **Bewachung**

65. *Context*: Der Polizist beschützte den Zeugen während des ganzen Prozesses.  
*Continuation*: Die Bewachung des Polizisten<sub>GenS</sub>/Zeugen<sub>GenO</sub> ließ die Familie in Ruhe schlafen.
66. *Context*: Der Agent schützte den Richter, bis der Prozess zu Ende war.  
*Continuation*: Der Minister hielt die Bewachung des Agenten<sub>GenS</sub>/Richters<sub>GenO</sub> für eine notwendige Maßnahme.

#### **Bewunderung**

67. *Context*: Der Enkel betete den Großvater an.  
*Continuation*: Die Bewunderung des Enkels<sub>GenS</sub>/Großvaters<sub>GenO</sub> machte den Vater manchmal etwas eifersüchtig.
68. *Context*: Das Mädchen vergötterte den Star bedingungslos.  
*Continuation*: Der Bruder hielt die Bewunderung des Mädchens<sub>GenS</sub>/Stars<sub>GenO</sub> für leicht verrückt.

#### **Diskriminierung**

69. *Context*: Der Lehrer benachteiligte einen Schüler konstant bei der Benotung.  
*Continuation*: Die Diskriminierung des Lehrers<sub>GenS</sub>/Schülers<sub>GenO</sub> zwang die Eltern zu einer Beschwerde.
70. *Context*: Der Beamte benachteiligte den älteren Bewerber bei der Vergabe des Postens.  
*Continuation*: Ein Kollege bemerkte die Diskriminierung des Beamten<sub>GenS</sub>/Bewerbers<sub>GenO</sub> von Anfang an.

#### **Durchleuchtung**

71. *Context*: Der Ermittler überprüfte den Verdächtigen wochenlang von Kopf bis Fuß.  
*Continuation*: Die Durchleuchtung des Ermittlers<sub>GenS</sub>/Verdächtigen<sub>GenO</sub> erbrachte den Beweis für eine Verurteilung.
72. *Context*: Der Staatsanwalt überprüfte den verdächtigen Politiker wieder und wieder.  
*Continuation*: Die Parteifreunde bezeichneten die Durchleuchtung des Staatsanwalts<sub>GenS</sub>/Politikers<sub>GenO</sub> als Hexenjagd.

#### **Durchsuchung**

73. *Context*: Der Polizist überprüfte den Räuber bei der Festnahme auf Waffen.  
*Continuation*: Die Durchsuchung des Polizisten<sub>GenS</sub>/Räubers<sub>GenO</sub> lockte einige Schaulustige in die Nähe.
74. *Context*: Der Zöllner kontrollierte bei der Einreise, ob der Fahrer Drogen bei sich hatte.  
*Continuation*: Ein Kollege beobachtete die Durchsuchung des Zöllners<sub>GenS</sub>/Fahrers<sub>GenO</sub> aus seinem Wagen heraus.

#### **Einlieferung**

75. *Context*: Der Notarzt brachte den Filmstar nach dem Zusammenbruch in die Klinik.  
*Continuation*: Die Einlieferung des Notarztes<sub>GenS</sub>/Filmstars<sub>GenO</sub> veranlasste die Presse zu wilden Spekulationen.
76. *Context*: Der Vater raste mit dem Mädchen nach ihrem Unfall in die Notaufnahme.  
*Continuation*: Ein Stau verzögerte die Einlieferung des Vaters<sub>GenS</sub>/Mädchens<sub>GenO</sub> um wenige Minuten.

#### Einschätzung

77. *Context*: Der Professor beschrieb den Studenten in seiner Empfehlung als brilliant.  
*Continuation*: Die Einschätzung des Professors<sub>GenS</sub>/Studenten<sub>GenO</sub> überzeugte den Vorsitzenden der Kommission.
78. *Context*: Der Experte prophezeite dem Team bei der WM ein Ausscheiden in der Vorrunde.  
*Continuation*: Die Fans bedachten die Einschätzung des Experten<sub>GenS</sub>/Teams<sub>GenO</sub> mit bösen Kommentaren.

#### Einschüchterung

79. *Context*: Der Gangster drohte dem Zeugen vor dem Prozess.  
*Continuation*: Die Einschüchterung des Gangsters<sub>GenS</sub>/Zeugen<sub>GenO</sub> zeigte keine Wirkung während der Aussage.
80. *Context*: Der Minister drohte dem Reporter mit einer Anzeige wegen Verleumdung.  
*Continuation*: Die Zeitung kommentierte die Einschüchterung des Ministers<sub>GenS</sub>/Reporters<sub>GenO</sub> in einer Stellungnahme.

#### Enteignung

81. *Context*: Der Richter beschlagnahmte das Vermögen des Betrügers für seine Opfer.  
*Continuation*: Die Enteignung des Richters<sub>GenS</sub>/Betrügers<sub>GenO</sub> gab den Opfern etwas Hoffnung.
82. *Context*: Das Ministerium verstaatlichte das Grundstück des Bauern für den Bau einer Straße.  
*Continuation*: Der Anwalt bekämpfte die Enteignung des Ministeriums<sub>GenS</sub>/Bauern<sub>GenO</sub> vor dem Höchstgericht.

#### Enterbung

83. *Context*: Der Vater strich den Sohn aus dem Testament.  
*Continuation*: Die Enterbung des Vaters<sub>GenS</sub>/Sohnes<sub>GenO</sub> stürzte die Familie in einen Streit.
84. *Context*: Der Großvater beschloss, den Enkel aus dem Testament zu nehmen.  
*Continuation*: Die Familie bekämpfte die Enterbung des Großvaters<sub>GenS</sub>/Enkels<sub>GenO</sub> vor Gericht.

#### Entführung

85. *Context*: Der Terrorist hielt den Diplomaten bis zur Lösegeldzahlung gefangen.  
*Continuation*: Die Entführung des Terroristen<sub>GenS</sub>/Diplomaten<sub>GenO</sub> hielt die Angehörigen in Atem.
86. *Context*: Der Täter verschleppte den Jungen in ein abgelegenes Versteck.  
*Continuation*: Die Familie bemerkte die Entführung des Täters<sub>GenS</sub>/Jungen<sub>GenO</sub> nach einigen Stunden.

#### Enthauptung

87. *Context*: Der Entführer schlug dem Gefangenen vor laufender Kamera den Kopf ab.  
*Continuation*: Die Enthauptung des Entführers<sub>GenS</sub>/Gefangenen<sub>GenO</sub> schockierte die Ermittler beim Ansehen des Videos.
88. *Context*: Der Henker führte den Verurteilten zum Schafott.  
*Continuation*: Hunderte Menschen verfolgten die Enthauptung des Henkers<sub>GenS</sub>/Verurteilten<sub>GenO</sub> mit großem Entsetzen.

#### Entlarvung

89. *Context*: Der Kommissar fand den wahren Mörder des Mädchens nach vielen Jahren.  
*Continuation*: Die Entlarvung des Kommissars<sub>GenS</sub>/Mörders<sub>GenO</sub> befreite den Vater von dem Mordverdacht.

90. *Context*: Der Reporter überführte den Radfahrer nach seinem Sieg des Dopings.  
*Continuation*: Die Polizei nutzte die Entlarvung des Reporters<sub>GenS</sub>/Radfahrers<sub>GenO</sub> für eine Razzia im Team-büro.

#### Entlassung

91. *Context*: Der Chirurg schickte den Jungen vier Monate nach der Operation nach Hause.  
*Continuation*: Die Entlassung des Chirurgen<sub>GenS</sub>/Jungen<sub>GenO</sub> freute das Personal auf der Station.
92. *Context*: Der Leiter des Büros warf den Angestellten ohne Grund aus der Firma.  
*Continuation*: Der Anwalt bekämpfte die Entlassung des Leiters<sub>GenS</sub>/Angestellten<sub>GenO</sub> vor Gericht.

#### Entlastung

93. *Context*: Der Assistent nahm dem Chef viel Arbeit ab.  
*Continuation*: Die Entlastung des Assistenten<sub>GenS</sub>/Chefs<sub>GenO</sub> erleichterte die Arbeit an dem neuen Projekt.
94. *Context*: Der Praktikant half dem Angestellten drei Monate lang mit seiner Arbeit.  
*Continuation*: Der Chef begrüßte die Entlastung des Praktikanten<sub>GenS</sub>/Angestellten<sub>GenO</sub> sehr.

#### Entmachtung

95. *Context*: Der Diktator entzog dem Parlament nach dem Putsch alle Kompetenzen.  
*Continuation*: Die Entmachtung des Diktators<sub>GenS</sub>/Parlaments<sub>GenO</sub> bedeutete das Ende für die Demokratie.
96. *Context*: Das Volk stürzte das Regime mit einer friedlichen Revolution.  
*Continuation*: Die Armee unterstützte die Entmachtung des Volkes<sub>GenS</sub>/Regimes<sub>GenO</sub> durch Untätigkeit.

#### Entwaffnung

97. *Context*: Der mutige Kunde riss dem Räuber die Pistole aus der Hand.  
*Continuation*: Die Entwaffnung des Kunden<sub>GenS</sub>/Räubers<sub>GenO</sub> zwang den Komplizen zu einer schnellen Flucht.
98. *Context*: Der Agent schlug dem Attentäter gerade noch rechtzeitig die Waffe aus der Hand.  
*Continuation*: Ein Reporter filmte die Entwaffnung des Agenten<sub>GenS</sub>/Attentäters<sub>GenO</sub> aus nächster Nähe.

#### Ergreifung

99. *Context*: Der Polizist schnappte den Mörder nach langen Ermittlungen.  
*Continuation*: Die Ergreifung des Polizisten<sub>GenS</sub>/Mörders<sub>GenO</sub> ermöglichte eine Verurteilung nach Jahren voller Rückschläge.
100. *Context*: Der Agent überwältigte den Terroristen nach einem kurzen Kampf.  
*Continuation*: Der Reporter schilderte die Ergreifung des Agenten<sub>GenS</sub>/Terroristen<sub>GenO</sub> in den Nachrichten.

#### Ermahnung

101. *Context*: Der Referee wies den Tormann nach einem Foul zurecht.  
*Continuation*: Die Ermahnung des Referees<sub>GenS</sub>/Tormanns<sub>GenO</sub> brachte die Fans zum Johlen.
102. *Context*: Der Trainer drohte dem Spieler, ihn bei weiteren Verfehlungen aus dem Kader zu nehmen.  
*Continuation*: Die Presse kommentierte die Ermahnung des Trainers<sub>GenS</sub>/Spielers<sub>GenO</sub> in zahllosen Artikeln.

#### Ermordung

103. *Context*: Der Attentäter tötete den Politiker während einer Wahlveranstaltung.  
*Continuation*: Die Ermordung des Attentäters<sub>GenS</sub>/Politikers<sub>GenO</sub> verzögerte die Wahl um einige Monate.
104. *Context*: Der Gangster erschoss den Richter, als er aus dem Haus kam.  
*Continuation*: Ein Zeuge schilderte die Ermordung des Gangsters<sub>GenS</sub>/Richters<sub>GenO</sub> mit zittriger Stimme.

#### Ermutigung

105. *Context*: Der Trainer richtete das Team nach der knappen Niederlage wieder etwas auf.  
*Continuation*: Die Ermutigung des Trainers<sub>GenS</sub>/Teams<sub>GenO</sub> machte die Niederlage für alle etwas erträglicher.

106. *Context:* Der Onkel bestärkte den Neffen in seinem Wunsch, zum Militär zu gehen.  
*Continuation:* Die Tante bezeichnete die Ermutigung des Onkels<sub>GenS</sub>/Neffen<sub>GenO</sub> als totalen Wahnsinn.

#### Erpressung

107. *Context:* Der ehemalige Assistent drohte dem Manager damit, belastende Daten weiterzuleiten.  
*Continuation:* Die Erpressung des Assistenten<sub>GenS</sub>/Managers<sub>GenO</sub> brachte die Konzernführung in eine schwierige Situation.
108. *Context:* Der Diener drohte dem Prinzen mit der Veröffentlichung von peinlichen Fotos.  
*Continuation:* Die Presse enthüllte die Erpressung des Dieners<sub>GenS</sub>/Prinzen<sub>GenO</sub> nach kurzer Zeit.

#### Förderung

109. *Context:* Der Millionär unterstützte den Künstler mit einer beträchtlichen Summe.  
*Continuation:* Die Förderung des Millionärs<sub>GenS</sub>/Künstlers<sub>GenO</sub> ermöglichte eine Arbeit ohne großen Druck.
110. *Context:* Der Konzern garantierte dem Verein aus der zweiten Liga Zahlungen über zwei Jahre.  
*Continuation:* Der Vorstand stoppte die Förderung des Konzerns<sub>GenS</sub>/Vereins<sub>GenO</sub> nach den zwei Jahren.

#### Freilassung

111. *Context:* Der Detektiv des Kaufhauses ließ den Jungen nach dem versuchten Diebstahl laufen.  
*Continuation:* Die Freilassung des Detektivs<sub>GenS</sub>/Jungen<sub>GenO</sub> ersparte der Familie eine Menge Ärger.
112. *Context:* Der Richter entließ den Entführer wegen Mangel an Beweisen aus der U-Haft.  
*Continuation:* Der Staatsanwalt kritisierte die Freilassung des Richters<sub>GenS</sub>/Entführers<sub>GenO</sub> in einer Stellungnahme.

#### Heilung

113. *Context:* Der Psychologe befreite den Patienten von seinen Angstzuständen.  
*Continuation:* Die Heilung des Psychologen<sub>GenS</sub>/Patienten<sub>GenO</sub> bedeutete einen Neubeginn für die ganze Familie.
114. *Context:* Der Arzt rettete den kranken Jungen mit einem neuen Medikament.  
*Continuation:* Das Medikament ermöglichte die Heilung des Arztes<sub>GenS</sub>/Jungen<sub>GenO</sub> ohne chirurgischen Eingriff.

#### Inhaftierung

115. *Context:* Der Inspektor steckte den Verdächtigen in Untersuchungshaft.  
*Continuation:* Die Inhaftierung des Inspektors<sub>GenS</sub>/Verdächtigen<sub>GenO</sub> brachte das Verfahren ins Rollen.
116. *Context:* Der Agent sollte den Terroristen in seiner Wohnung festnehmen.  
*Continuation:* Ein Fehler verhinderte die Inhaftierung des Agenten<sub>GenS</sub>/Terroristen<sub>GenO</sub> in letzter Minute.

#### Manipulierung

117. *Context:* Der Pate warnte den Zeugen vor einer Aussage gegen ihn.  
*Continuation:* Die Manipulierung des Paten<sub>GenS</sub>/Zeugen<sub>GenO</sub> gab dem Prozess eine unerwartete Wende.
118. *Context:* Der Gangster versprach dem Geschworenen im Fall eines Freispruchs eine hohe Summe.  
*Continuation:* Eine Reporterin enthüllte die Manipulierung des Gangsters<sub>GenS</sub>/Geschworenen<sub>GenO</sub> nach dem Ende des Prozesses.

#### Rehabilitierung

119. *Context:* Der Richter hob das Urteil gegen den Mann nach 12 Jahren auf.  
*Continuation:* Die Rehabilitierung des Richters<sub>GenS</sub>/Mannes<sub>GenO</sub> rührte den Anwalt zu Tränen.
120. *Context:* Der Detektiv bewies die Unschuld des Angeklagten vor Gericht.  
*Continuation:* Die Presse kommentierte die Rehabilitierung des Detektivs<sub>GenS</sub>/Angeklagten<sub>GenO</sub> in zahlreichen Berichten.

### Rettung

121. *Context:* Der Passant holte den Fahrer in letzter Minute aus dem Autowrack.  
*Continuation:* Die Rettung des Passanten<sub>GenS</sub>/Fahrers<sub>GenO</sub> beherrschte die Titelblätter in den lokalen Zeitungen.
122. *Context:* Der Hund fand das Opfer der Lawine noch rechtzeitig unter dem Schnee.  
*Continuation:* Der Einsatzleiter beobachtete die Rettung des Hundes<sub>GenS</sub>/Opfers<sub>GenO</sub> aus dem Helikopter.

### Schulung

123. *Context:* Der Experte erklärte dem Arbeiter in dem Seminar die neue Maschine.  
*Continuation:* Die Schulung des Experten<sub>GenS</sub>/Arbeiter<sub>GenO</sub> erleichterte die Arbeit an der Maschine um Einiges.
124. *Context:* Der Experte sollte das Team an einem Abend über Sicherheit am Arbeitsplatz aufklären.  
*Continuation:* Ein Vorfall verzögerte die Schulung des Experten<sub>GenS</sub>/Teams<sub>GenO</sub> um einige Tage.

### Täuschung

125. *Context:* Der Verkäufer führte den Jungen beim Kauf hinters Licht.  
*Continuation:* Die Täuschung des Verkäufers<sub>GenS</sub>/Jungen<sub>GenO</sub> verärgerte den Vater ungemein.
126. *Context:* Der Betrüger legte den Gutachter der Versicherung herein.  
*Continuation:* Ein Detektiv enthüllte die Täuschung des Betrügers<sub>GenS</sub>/Gutachters<sub>GenO</sub> nach einiger Zeit.

### Tötung

127. *Context:* Der Agent erschoss den Attentäter nach dem mißglückten Anschlag.  
*Continuation:* Die Tötung des Agenten<sub>GenS</sub>/Attentäters<sub>GenO</sub> verursachte große Panik unter den Anwesenden.
128. *Context:* Der Wilderer erschoss den Löwen mitten im Nationalpark.  
*Continuation:* Ein Ranger beobachtete die Tötung des Wilderers<sub>GenS</sub>/Löwen<sub>GenO</sub> aus der Ferne.

### Überprüfung

129. *Context:* Der Beamte prüfte den Hintergrund des Touristen vor der Einreise per Computer.  
*Continuation:* Die Überprüfung des Beamten<sub>GenS</sub>/Touristen<sub>GenO</sub> verzögerte die Einreise um einige Minuten.
130. *Context:* Der Detektiv durchleuchtete den Verdächtigen von Kopf bis Fuß.  
*Continuation:* Der Auftraggeber bezahlte die Überprüfung des Detektivs<sub>GenS</sub>/Verdächtigen<sub>GenO</sub> im Voraus.

### Überstellung

131. *Context:* Der Wärter begleitete den Häftling in ein neues Gefängnis.  
*Continuation:* Die Überstellung des Wärters<sub>GenS</sub>/Häftlings<sub>GenO</sub> ermöglichte eine Verwahrung ohne Risiko.
132. *Context:* Der Pfleger brachte den Patienten von der Intensivstation in ein normales Zimmer.  
*Continuation:* Der Arzt bestätigte die Überstellung des Pflegers<sub>GenS</sub>/Patienten<sub>GenO</sub> mit seiner Unterschrift.

### Überwachung

133. *Context:* Der Agent folgte dem Diplomaten über mehrere Wochen.  
*Continuation:* Die Überwachung des Agenten<sub>GenS</sub>/Diplomaten<sub>GenO</sub> erbrachte keine Hinweise auf mögliche Spionage.
134. *Context:* Der Detektiv beobachtete den Mann über mehrere Monate.  
*Continuation:* Der Auftraggeber finanzierte die Überwachung des Detektivs<sub>GenS</sub>/Mannes<sub>GenO</sub> mit einem Kredit.

### Umarmung

135. *Context:* Der Vater schloss den Sohn nach 20 Jahren Streit in seine Arme.  
*Continuation:* Die Umarmung des Vaters<sub>GenS</sub>/Sohnes<sub>GenO</sub> rührte die Mutter zu Tränen.

136. *Context:* Der Papst schloss den Pilger in seine Arme.  
*Continuation:* Der Regisseur zeigte die Umarmung des Papstes<sub>GenS</sub>/Pilgers<sub>GenO</sub> in einer Nahaufnahme.

#### Unterdrückung

137. *Context:* Der Tyrann unterwarf das Volk mit brutalen Mitteln.  
*Continuation:* Die Unterdrückung des Tyrannen<sub>GenS</sub>/Volkes<sub>GenO</sub> forderte viele Opfer unter den Bürgern.
138. *Context:* Der Diktator beutete das Volk brutal und gnadenlos aus.  
*Continuation:* Ein Putsch beendete die Unterdrückung des Diktators<sub>GenS</sub>/Volkes<sub>GenO</sub> nach drei Jahren.

#### Unterstützung

139. *Context:* Der Konzern förderte den Sportler mit einer großen Summe.  
*Continuation:* Die Unterstützung des Konzerns<sub>GenS</sub>/Sportlers<sub>GenO</sub> hatte positive Auswirkungen auf das Image der Firma.
140. *Context:* Der Senat bewilligte dem umstrittenen Künstler einen großen Zuschuss für seine Aktion.  
*Continuation:* Eine Zeitung kritisierte die Unterstützung des Senats<sub>GenS</sub>/Künstlers<sub>GenO</sub> auf dem Titelblatt.

#### Untersuchung

141. *Context:* Der Doktor überprüfte den Mann auf mögliche neurologische Störungen.  
*Continuation:* Die Untersuchung des Doktors<sub>GenS</sub>/Mannes<sub>GenO</sub> brachte keinen Hinweis auf eine Erkrankung.
142. *Context:* Der Ermittler der Steuerbehörde überprüfte den Unternehmer nach einem Hinweis.  
*Continuation:* Der Vorgesetzte genehmigte die Untersuchung des Ermittlers<sub>GenS</sub>/Unternehmers<sub>GenO</sub> mit seiner Unterschrift.

#### Verachtung

143. *Context:* Der Soldat hasste den Offizier von Anfang an.  
*Continuation:* Die Verachtung des Soldaten<sub>GenS</sub>/Offiziers<sub>GenO</sub> belastete die Stimmung in der Truppe.
144. *Context:* Der Mann empfand eine tiefe Abscheu gegen den Kanzler.  
*Continuation:* Ein Freund bezeichnete die Verachtung des Mannes<sub>GenS</sub>/Kanzlers<sub>GenO</sub> als nahezu pathologisch.

#### Verbannung

145. *Context:* Der König jagte den aufständischen General aus dem Land.  
*Continuation:* Die Verbannung des Königs<sub>GenS</sub>/Generals<sub>GenO</sub> beendete den Aufstand gegen die Monarchie.
146. *Context:* Der Bischof schickte den ketzerischen Mönch ins Exil.  
*Continuation:* Eine Urkunde belegt die Verbannung des Bischofs<sub>GenS</sub>/Mönchs<sub>GenO</sub> in allen Details.

#### Verehrung

147. *Context:* Das Mädchen himmelte den Musiker so sehr an, dass sie zu jedem Konzert reiste.  
*Continuation:* Die Verehrung des Mädchens<sub>GenS</sub>/Musikers<sub>GenO</sub> beunruhigte die Eltern zusehends.
148. *Context:* Der Fan vergötterte das Team sogar noch, als es in die vierte Liga abgestiegen war.  
*Continuation:* Seine Freunde bezeichneten die Verehrung des Fans<sub>GenS</sub>/Teams<sub>GenO</sub> als verrückt.

#### Verhaftung

149. *Context:* Der Agent fasste den Terroristen nach einem Tipp aus der Bevölkerung.  
*Continuation:* Die Verhaftung des Agenten<sub>GenS</sub>/Terroristen<sub>GenO</sub> verhinderte einen Anschlag auf einen Zug.
150. *Context:* Der Polizist schnappte den Mörder nach einer wilden Verfolgungsjagd.  
*Continuation:* Der Reporter filmte die Verhaftung des Polizisten<sub>GenS</sub>/Mörders<sub>GenO</sub> für die Nachrichten.

#### Verlegung

151. *Context:* Der Pfleger brachte den Patienten von der Intensivstation in die Chirurgie.  
*Continuation:* Die Verlegung des Pflegers<sub>GenS</sub>/Patienten<sub>GenO</sub> überraschte den Chefarzt der Chirurgie.



152. *Context:* Der Wächter wollte den Gefangenen in einen anderen Trakt bringen.  
*Continuation:* Der Direktor bewilligte die Verlegung des Wächters<sub>GenS</sub>/Gefangenen<sub>GenO</sub> nach einem Gespräch.

#### **Vernehmung**

153. *Context:* Der Kommissar befragte den verhafteten Kriminellen über die Organisation.  
*Continuation:* Die Vernehmung des Kommissars<sub>GenS</sub>/Kriminellen<sub>GenO</sub> brachte viele Informationen für den Staatsanwalt.
154. *Context:* Der Ankläger befragte den Zeugen zu Beginn des Prozesses.  
*Continuation:* Der Verteidiger störte die Vernehmung des Anklägers<sub>GenS</sub>/Zeugen<sub>GenO</sub> mit vielen Einsprüchen.

#### **Verschleppung**

155. *Context:* Der Spion entführte den Forscher aus seinem Heimatland.  
*Continuation:* Die Verschleppung des Spions<sub>GenS</sub>/Forschers<sub>GenO</sub> verschärfte die Spannungen zwischen den Ländern.
156. *Context:* Der Vater brachte das Kind ohne Wissen der Mutter in sein Heimatland.  
*Continuation:* Die Mutter meldete die Verschleppung des Vaters<sub>GenS</sub>/Kindes<sub>GenO</sub> an die Polizei.

#### **Versetzung**

157. *Context:* Der General wies den Soldaten einer neuen Einheit zu.  
*Continuation:* Die Versetzung des Generals<sub>GenS</sub>/Soldaten<sub>GenO</sub> überraschte die Kollegen in der Kompanie.
158. *Context:* Der Chef schickte den Mitarbeiter zum Aufbau des neuen Büros nach Asien.  
*Continuation:* Die Kollegen bedauerten die Versetzung des Chefs<sub>GenS</sub>/Mitarbeiters<sub>GenO</sub> ziemlich.

#### **Versorgung**

159. *Context:* Der Sanitäter leistete dem Opfer Erste Hilfe.  
*Continuation:* Die Versorgung des Sanitäters<sub>GenS</sub>/Opfers<sub>GenO</sub> ermöglichte den Transport in das nächste Krankenhaus.
160. *Context:* Der Koch stellte während der Expedition sicher, dass das Team gut gepflegt wurde.  
*Continuation:* Der Leiter lobte die Versorgung des Kochs<sub>GenS</sub>/Teams<sub>GenO</sub> nach dem Abenteuer.

#### **Verteidigung**

161. *Context:* Der Anwalt vertrat den Manager in einem schwierigen Prozess.  
*Continuation:* Die Verteidigung des Anwalts<sub>GenS</sub>/Managers<sub>GenO</sub> erforderte eine Strategie mit viel Risiko.
162. *Context:* Das Team hielt den Stürmer das ganze Spiel über unter Kontrolle.  
*Continuation:* Der Trainer lobte die Verteidigung des Teams<sub>GenS</sub>/Stürmers<sub>GenO</sub> nach dem Spiel.

#### **Vertreibung**

163. *Context:* Der Besitzer zwang den Mieter, aus der Wohnung auszuziehen.  
*Continuation:* Die Vertreibung des Besitzers<sub>GenS</sub>/Mieters<sub>GenO</sub> veranlasste den Anwalt zu einer Klage.
164. *Context:* Der General jagte den König nach dem Putsch aus dem Land.  
*Continuation:* Die UNO verurteilte die Vertreibung des Generals<sub>GenS</sub>/Königs<sub>GenO</sub> in einer Stellungnahme.

#### **Verurteilung**

165. *Context:* Der Richter schickte den Mann für zehn Jahre ins Gefängnis.  
*Continuation:* Die Verurteilung des Richters<sub>GenS</sub>/Mannes<sub>GenO</sub> überraschte die Zuseher im Gerichtssaal.
166. *Context:* Das Gremium sperrte den Sportler nach dem positiven Dopingtests für zwei Jahre.  
*Continuation:* Der Verband verkündete die Verurteilung des Gremiums<sub>GenS</sub>/Sportlers<sub>GenO</sub> in einer Pressemitteilung.

#### **Würdigung**

167. *Context:* Der Vorgesetzte lobte den Angestellten in der Besprechung in den Himmel.  
*Continuation:* Die Würdigung des Vorgesetzten<sub>GenS</sub>/Angestellten<sub>GenO</sub> beeindruckte die Kollegen in der Firma.
168. *Context:* Der Minister betonte in seiner Rede die Verdienste des Künstlers.  
*Continuation:* Das Fernsehen zeigte die Würdigung des Ministers<sub>GenS</sub>/Künstlers<sub>GenO</sub> in einem Beitrag.